

TECHNISCHE UNIVERSITÄT BERLIN
FAKULTÄT II - MATHEMATIK UND NATURWISSENSCHAFTEN
INSTITUT FÜR MATHEMATIK

**Ein mathematisches Modell zur
Verkehrszustandsschätzung mit Bayes'schen Netzen
- Geschwindigkeitsschätzung durch Fusion von
C2X- und Bluetooth Daten**

Masterarbeit

zur Erlangung des akademischen Grades
Master of Science (M. Sc.)



eingereicht von: Kim Klüber
geboren am: 11.02.1993
geboren in: Bad Hersfeld
Matrikelnummer: 378122

Gutachter: Priv.-Doz. Dr. rer. nat. Konstantin Fackeldey
Dr.-Ing. Marek Junghans

Eidesstattliche Versicherung

Die selbständige und eigenhändige Anfertigung versichert an Eides statt

Berlin, den 4. Januar 2018

.....
Unterschrift

Inhaltsverzeichnis

1	Motivation	5
2	Stand von Wissenschaft und Technik	10
2.1	Verkehrliche Grundlagen	10
2.1.1	Verkehrszustandsbestimmung	10
2.1.2	Lichtsignalanlagen	12
2.2	Kommunikation und Sensoren	14
2.2.1	Car-2-X-Kommunikation	14
2.2.2	Bluetooth	17
2.3	Verkehrssimulator SUMO	19
2.3.1	SUMO	19
2.3.2	Fahrzeugfolgemodell von Krauß	20
2.3.3	Szenarien	22
2.3.4	TraCI	28
3	Bayes'sche Wahrscheinlichkeit	30
3.1	Bayes'sche und klassische Wahrscheinlichkeitstheorie	30
3.2	Bayes'sche Netze	34
3.3	Wahrscheinlichkeiten in Bayes'schen Netzen	36
3.4	Datenfusion	39
3.5	Zugrunde liegendes Modell	41
4	Modell und Implementierung	47
4.1	Modellierung des Bayes'schen Netzes	47
4.1.1	Knotenwahl und kausale Zusammenhänge	48
4.1.2	Diskretisierung der Knoten	53
4.2	Implementierung	57
4.2.1	Detektion	58
4.2.2	Feststellen der relativen Häufigkeiten	60
4.2.3	Schätzen und Fusion der Daten	62
5	Simulative Untersuchungen	65
5.1	Aufbau der Untersuchungen	65
5.2	Ergebnisse	69
5.2.1	Szenario Minimal	70

5.2.2	Szenario RiLSA	76
5.2.3	Szenario BS Forschungskreuzung	82
6	Fazit	85

1 Motivation

Ein vorrangiges gesellschaftliches Ziel ist, die Mobilität von Personen und Gütern zu verbessern. Dabei ist der Straßenverkehr von großer Bedeutung. Aktuell steigen die Ansprüche an die Mobilität immer weiter. Um diesen gerecht zu werden, müssen neue Technologien für das Verkehrsmanagement nutzbar gemacht werden. Dabei sind neue Informationstechnologien zugleich Chance und Herausforderung. Die fortschreitende Automatisierung von Fahrzeugen bietet die vielversprechende Möglichkeit, das Verkehrsmanagement durch Sensordaten und Nachrichtenaustausch direkt zu unterstützen. Hierfür müssen neue Methoden entwickelt werden und es muss definiert werden, wie Verkehr künftig funktionieren soll. [32]

Eine neue Kommunikationstechnologie zum Austausch kooperativer Verkehrsdaten, die *Car-2-X-Kommunikation* - kurz C2X -, kommt auf den Straßen zum Einsatz. Sie umfasst die Kommunikation zwischen zwei Fahrzeugen einerseits und die zwischen einem Fahrzeug und bestehender Infrastruktur andererseits. Sie findet im automatisierten Fahren schon heute Anwendung und soll in Zukunft die Zusammenführung von Sensordaten vieler Fahrzeuge ermöglichen. Mit C2X soll ein neues Level an Verkehrssicherheit und -effizienz geschaffen werden. [4], [32]

In ersten Untersuchungen des Potentials der C2X-Kommunikation auf Autobahnen wurde die Kommunikationstechnologie genutzt, um Verkehrsinformationen wie etwa Staumeldungen zu verbreiten. Ein Fahrzeug, das in den betroffenen Bereich einfährt, detektiert den Stau und sendet die Staumeldung weiter an umliegende Fahrzeuge. Diese transportieren die Meldung wieder weiter an alle Fahrzeuge in Reichweite. So breitet sich die Information immer weiter aus. Obwohl die Ausstattungsrate der C2X-fähigen Fahrzeuge nur bei 2% lag, konnten so die Verkehrsinformationen über 50 km verbreitet werden. [31]

Eine weitere Analyse zur Verbreitung von Unfallmeldungen zeigte den Nutzen von C2X im innerstädtischen Bereich. Die Meldungen über einen Unfall, der sich im Stadtkern ereignet hatte, wurden dabei von den mit C2X ausgestatteten Fahrzeugen weitergeleitet. Zusätzlich wurden für die Verteilung der Information auch Infrastrukturstützstellen verwendet. An diesen Stellen wurden die Meldungen lokal gespeichert und stets weiterverteilt. Dadurch konnte die Meldung auch Fahrzeuge erreichen, die sich erst zu einem späteren Zeitpunkt in Kommunikationsreichweite befanden. Insgesamt erreichten die Meldungen nach nur 500 s 90% der ausgestatteten Fahrzeuge im Umkreis von 3 km. [47]

Aktuell sind nur vereinzelt Fahrzeuge mit C2X-Technologie ausgestattet. Um Daten vieler Fahrzeuge zu verknüpfen, muss zunächst ein höherer Ausstattungsgrad erreicht werden. Aufgrund der intensiven Erforschung und Weiterentwicklung der C2X-Technologie

wird prognostiziert, dass ihr Anteil auf den Straßen weiter zunimmt. Diese Entwicklungen und Voraussagen sind derzeit noch spekulativ. Treten diese Prognosen ein und automatisierte Fahrzeuge übernehmen Stück für Stück Oberhand auf den Straßen, so ist der Normalzustand ein Mischverkehr aus automatisierten und nicht-automatisierten Fahrzeugen. Effekte dieses Verkehrs müssen untersucht und bestehende Instrumente angepasst werden, um alle Verkehrsteilnehmer im Verkehrssystem zu berücksichtigen. [45]

Die vorliegende Arbeit befasst sich mit der Verkehrszustandsschätzung an Kreuzungen mit Lichtsignalanlage, umgangssprachlich Ampel genannt. Dabei wird C2X-Kommunikation genutzt. Sogenannte Roadside Units werden entlang der Straße aufgebaut und mit der für die Kommunikation erforderlichen Technik ausgestattet. Sie ermöglichen die Kommunikation zwischen Fahrzeugen und Infrastruktur. Betrachtet wird der Fall, dass ein Mischverkehr aus Fahrzeugen mit und ohne C2X-Ausstattung vorliegt.

In der Arbeit soll untersucht werden, inwieweit sich die C2X-Kommunikation für die Verkehrszustandsschätzung nutzen lässt. Besonders der Einfluss der Ausstattungsrate der Fahrzeuge mit C2X soll dabei analysiert werden.

Wenn sich die Ausstattungsrate in einem niedrigen Bereich $< 20\%$ bewegt, können nur wenige Daten erfasst werden. Dies ist problematisch, da so nur einzelne Fahrzeuge erfassbar sind und die Detektion des Verkehrs insgesamt unvollständig bleibt. Dieses Problem kann behoben werden, indem ein weiterer Sensor genutzt wird, um zusätzliche Daten zu erfassen. Mit einem Bluetooth-basierten Verfahren werden weitere Fahrzeuge detektiert und deren Geschwindigkeiten gemessen. Da viele Verkehrsteilnehmer Bluetoothgeräte wie Navigationsgeräte, Handys oder Freisprecheinrichtungen mitführen, ist davon auszugehen, dass die Bluetooth-Penetrationsrate höher ist als die der C2X-Sender. Als realistisch werden Ausstattungsgrade von 30-50% für Bluetooth angenommen [37]. Experimente werden mit 1 - 100% C2X-Penetration durchgeführt.

Mit C2X können GPS-Positionen und daraus resultierende Geschwindigkeiten übermittelt werden. Das GPS-Verfahren zur Positionsbestimmung liefert auf freier Fläche auf etwa 5 m genaue Ergebnisse. In der Nähe von Störfaktoren wie Gebäuden oder Bäumen verschlechtert sich die Genauigkeit. Unter zusätzlichem Zugriff auf Wide Area Augmentation System-Satelliten kann die Genauigkeit auf weniger als 3 m verbessert werden. [16], [18]

Diese Geschwindigkeitsdaten sind daher in der Regel genauer als die Messwerte aus dem Bluetooth-Verfahren. Dies bringt Schwankungen der Detektionszeitpunkte mit sich, die zu fehlerbehafteten Geschwindigkeitsschätzungen führen. Die Bluetooth-Detektion ermöglicht nur eine sehr vage Geschwindigkeitsschätzung. Somit stehen einerseits genauere C2X-Daten in geringer Menge und andererseits Bluetooth-Daten in größerer Menge zur Verfügung. Da die Funktionsweise von Bluetooth Schwankungen der Detektionszeitpunkte mit sich bringt, sind die entstehenden Daten fehlerbehafteter als die C2X-Daten. Ein Trade-Off zwischen Datenmenge und Datenqualität besteht.

Sind Daten aus unterschiedlichen Quellen verfügbar, kann eine Datenfusion durch-

geführt werden. Diese bezeichnet das Zusammenführen von Daten unterschiedlichen Ursprungs zu gemeinsamen Informationen zu einer bestimmten Fragestellung. Oft kann so die Qualität der einzelnen Daten übertroffen werden. [19]

Ziel der Arbeit ist, Aussagen über den Verkehr an einem Knotenpunkt zu treffen. Mittels Datenfusion soll die mittlere Momentangeschwindigkeit bestimmt werden. Diese gibt Aufschluss über die Verkehrsflüsse der jeweiligen Richtungen. Das Modell stellt den Zusammenhang zwischen tatsächlicher Geschwindigkeit und detektierter Bluetooth- und C2X-Geschwindigkeit dar. Weitere Faktoren, wie wichtige Verkehrskenngrößen, werden vorgestellt, auf ihre Relevanz hin untersucht und entsprechend integriert.

Zur Modellierung wird ein Ansatz basierend auf Bayes'schen Netzen verwendet. Diese sind gerichtete Graphen, die aus Knoten und Kanten bestehen und zur Modellierung kausaler Zusammenhänge genutzt werden können. Dabei repräsentieren die Knoten Zufallsvariablen, die jedem möglichen Ergebnis des Experiments Wahrscheinlichkeiten zuordnen. Daher müssen die Graphen zusätzlich mit Wahrscheinlichkeitsverteilungen versehen sein.

Bayes'sche Netze bieten den Vorteil, dass es zunächst ausreicht, relevante Faktoren zu identifizieren und Kausalitäten festzustellen. Sind die Knoten nach einem Ursache-Wirkungs-Prinzip verbunden, so kann das Bayes'sche Netz anhand von Daten aus Experimenten mit Wahrscheinlichkeiten quantifiziert werden. Eine analytische Modellierung der einzelnen Ursache-Wirkungs-Prinzipien ist dann nicht erforderlich.

In Abbildung 1.1 ist ein Bayes'sches Netz dargestellt. Dabei sei X die unbekannte Zielgröße und die grau hinterlegten Knoten U_1 , U_2 , W_1 und W_2 Beobachtungen, die getroffen werden können. Weiter sei eine Wahrscheinlichkeitsverteilung gegeben.

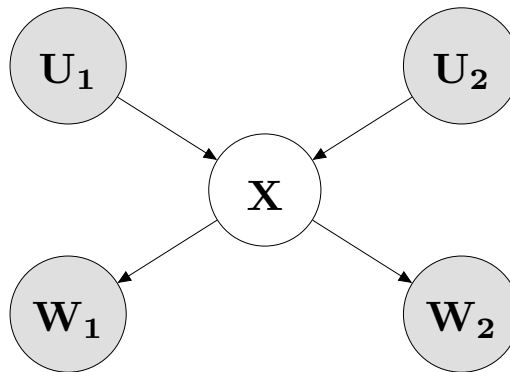


Abbildung 1.1: Beispiel für einen gerichteten Graphen, der zusammen mit der Wahrscheinlichkeitsverteilung P ein Bayes'sches Netz beschreibt. U_1 , U_2 , W_1 und W_2 sind Beobachtungen, X die unbekannte Zielgröße. Die gerichteten Kanten stellen Ursache-Wirkungs-Beziehungen dar.

Bayes'sche Netze können dazu verwendet werden, von Beobachtungen Rückschlüsse auf deren unbekannte Ursachen oder Wirkungen zu ziehen. Ist in Abbildung 1.1 die Größe X unbekannt, während U_1 , U_2 , W_1 und W_2 bekannt sind, können letztere genutzt werden, um X zu schätzen.

Dabei wird zwischen diagnostischer und kausaler Unterstützung unterschieden.

1. **Kausale Unterstützung** beschreibt Beobachtungen der Ursachen eines Phänomens. Es lässt darauf schließen, dass deren Wirkung wahrscheinlich ist. U_1 und U_2 bieten kausale Unterstützung für X .
2. **Diagnostische Unterstützung** beschreibt hingegen die Beobachtung einer Folge, die auf deren Ursache schließen lässt. Im Beispiel unterstützen W_1 und W_2 die Unbekannte X diagnostisch. Sie können als Wirkung von X beobachtet werden.

Mit diagnostischer und kausaler Unterstützung kann die Zielgröße unter Ausnutzung des Satzes von Bayes berechnet werden [35].

Der Ansatz mit Bayes'schen Netzen vereinfacht es, Fachwissen und Daten zu kombinieren. Zusammenhänge, die analytisch schwer beschreibbar sind, können zunächst qualitativ in kausale Beziehung gebracht und schließlich experimentell quantifiziert werden.

Daher soll im Folgenden ein Bayes'sches Netz aufgestellt werden, das eine Geschwindigkeitsschätzung auf Grundlage der C2X- und Bluetooth-Informationen ermöglicht. Die Kombination von Bluetooth und C2X-Daten zur Verkehrszustandsschätzung sowie der Ansatz basierend auf Bayes'schen Netzen sind aus [24] übernommen und werden weiterentwickelt.

In der vorliegenden Arbeit sollen C2X- und Bluetooth-Detektion in Relation zur tatsächlichen Geschwindigkeit gesetzt und relevante Einflussfaktoren identifiziert werden. Diese Größen werden als Zufallsvariablen interpretiert und anschließend wird ein mathematisches Modell basierend auf einem Bayes'schen Netz aufgestellt. Zur Datengenerierung werden simulative Experimente im Verkehrssimulator SUMO durchgeführt. Das Netz kann so mit relativen Häufigkeiten quantifiziert werden. Implementierungen der Simulationsläufe und Datenerfassungen sowie ein Skript zur Berechnung der relativen Häufigkeiten und Speicherung der multidimensionalen Matrizen werden erstellt. Schließlich wird die modellierte Datenfusion in Echtzeit implementiert, die in erneuten Simulationsläufen integriert und getestet werden kann. Anschließend werden die Ergebnisse mit den echten Ergebnissen im Hinblick auf Genauigkeit und Vollständigkeit verglichen. Insbesondere wird der Einfluss der Ausstattungsgrade mit Bluetooth und C2X-Kommunikation auf die Geschwindigkeitsschätzung untersucht.

Im folgenden Kapitel 2 wird das zur Rede stehende Problem anhand des Stands der Technik analysiert. Es werden wichtige Voraussetzungen für die Untersuchungen auf den Gebieten verkehrlicher Grundlagen, Kommunikation und Sensoren und Verkehrssimulation mit SUMO vorgestellt. Kapitel 3 ist der Bayes'schen Wahrscheinlichkeitstheorie

und Bayes'schen Netzen im Speziellen gewidmet. Abschließend wird die spezielle Aufgabenstellung konkret definiert. Das Kapitel 4 beinhaltet die mathematische Modellierung der Datenfusion basierend auf den zuvor erklärten Methoden. Die Implementierung des Modells und dessen Anwendung werden erklärt. Das Kapitel 5 enthält den Simulationsaufbau und die Ergebnisse der simulativen Untersuchungen. Im letzten Kapitel 6 werden die Ergebnisse zusammengefasst und ein Ausblick für weiterführende Untersuchungen gegeben.

2 Stand von Wissenschaft und Technik

Ziel der Arbeit ist, den Verkehrszustand an einem Knotenpunkt zu erfassen. Um ein aussagekräftiges Modell zu erstellen, sollen in diesem Kapitel zunächst Teilaspekte der Untersuchung näher beleuchtet werden. So werden verkehrliche Grundlagen gelegt, um die Frage zu klären, wie ein Verkehrszustand messbar gemacht werden kann. Damit können verkehrliche Kenngrößen und nutzbare Daten identifiziert werden, die im Modell betrachtet werden können. Weiter werden Zusammenhänge zwischen den Kenngrößen thematisiert, um einen Rahmen für Untersuchungen der Kausalitäten zu liefern.

Die Verkehrszustandsbestimmung soll unter Rückgriff auf C2X-Kommunikation und Bluetooth-Detektion erfolgen. In diesem Kapitel wird erklärt, wie Bluetooth-Detektion und C2X-Kommunikation funktionieren und wie Verfahren gewählt werden können, die aus einer Bluetooth-Detektion eine Geschwindigkeit ableitbar machen. Es wird auch erklärt, wie C2X-Kommunikation genutzt werden kann. Mit diesen Verfahren soll die Grundlage für die Datenfusion geschaffen werden. Eine erste Analyse, die zeigt, welche Ergebnisse möglich sind, schließt den Abschnitt ab.

Anschließend folgt eine Beschreibung des Verkehrssimulators SUMO, der für simulative Experimente genutzt werden soll. Dies umfasst eine Vorstellung der simulativ untersuchten Szenarien, des zugrundeliegenden Fahrzeugfolgmodells und des Interfaces TraCI - kurz für *Traffic Control Interface* - zur Steuerung der Verkehrssimulation.

2.1 Verkehrliche Grundlagen

In diesem Abschnitt werden für das Verkehrsmanagement relevante Zusammenhänge und Größen vorgestellt. Zunächst wird erklärt, was eine Verkehrszustandsschätzung ist und wofür sie genutzt wird. Zugrundeliegende Kenngrößen werden eingeführt. Dann werden Lichtsignalanlagen beschrieben und ihre Funktionsweise erläutert.

2.1.1 Verkehrszustandsbestimmung

Um ein effizientes Verkehrsmanagement zu gewährleisten, müssen Verkehrszustände erfassbar und prognostizierbar sein. Informationen über den Verkehrszustand können dann für Verkehrsteilnehmer bereitgestellt oder zur Verkehrssteuerung genutzt werden. Das Feststellen des Verkehrszustands wird als Verkehrszustandsbestimmung bezeichnet. [39]

Der Verkehrsablauf ist durch Individualität, Kollektivität und Komplexität charakterisiert. Das Verhalten aller Verkehrsteilnehmer ist individuell, obgleich sie sich als Kollektiv in ihrem Fahrverhalten ähneln. Der Verkehr folgt dabei komplexen Strukturen, die zwar durch Verkehrsregeln gelenkt, aufgrund subjektiver Vorlieben und Einschätzungen aller Verkehrsteilnehmer jedoch schwer in Gänze erfassbar sind.

Im Straßenverkehr auftretende Erscheinungen wie etwa Staubbildung oder Beschleunigung und Abbremsung an Knotenpunkten lassen sich dabei mathematisch-statistisch beschreiben und erforschen.

Die Modellierung des Verkehrs erlaubt es, ihn auch simulativ zu untersuchen. So können spezielle Szenarien rekonstruiert und erforscht werden. Die verwendeten Modelle werden in deterministische und stochastische Modelle unterteilt. Während deterministische Modelle die Zufälligkeit des Verkehrsflusses vernachlässigen und den Verkehrsteilnehmern ein gleichartiges mittleres Verhalten zuordnen, bilden stochastische Modelle diese Zufälligkeit ab und sind so aussagekräftiger. Oft ist ihre Modellierung mathematisch schwieriger, weshalb auf Simulationen zurückgegriffen wird. [41]

Eine weitere Unterscheidung der Modelle besteht in ihrem Fokus. Mikroskopische Modelle erfassen die Bewegung des Einzelfahrzeugs, während makroskopische Modelle den Verkehr in seiner Gesamtheit beschreiben und dessen stochastische Eigenschaften berücksichtigen.

Um ein solches Modell zur Verkehrszustandsschätzung zu erstellen, müssen verkehrszustandsrelevante Größen identifiziert und mittels geeigneter Detektionseinrichtungen erfassbar gemacht werden. [39]

Nach Schnabel und Lohse [41] sind die wichtigsten Verkehrskenngrößen:

- Verkehrsstärke oder Intensität,
- Dichte oder Konzentration,
- Geschwindigkeit,
- zeitliche Abstände der Fahrzeuge untereinander,
- räumliche Abstände zwischen den Fahrzeugen.

Dabei bezeichnet die **Verkehrsstärke** q den Quotienten aus der Fahrzeuganzahl N über einer Zeitspanne T . Es gilt:

$$q = N/T. \quad (2.1)$$

Die Verkehrsstärke q wird meist in Fahrzeugen pro Stunde angegeben. Wird die Verkehrsstärke in Sekunden angegeben, wird sie auch als Intensität bezeichnet.

Die **Verkehrsdichte** D ist die Anzahl der Fahrzeuge über einem Wegabschnitt der Länge S :

$$D = N/S. \quad (2.2)$$

In der Regel wird ein Wegabschnitt von 1 km als Bezugsgröße verwendet.

Die Geschwindigkeit v wird in m/s oder km/h gemessen und kann lokal oder momentan betrachtet werden.

Die **lokale Geschwindigkeit** wird an einem festen Querschnitt für eine festgelegte Zeitdauer ermittelt. Dann ist die mittlere lokale Geschwindigkeit \bar{v}_l für N Messwerte und die lokale Geschwindigkeit v_{li} eines Fahrzeugs i

$$\bar{v}_l = \frac{1}{N} \sum_{i=1}^N v_{li}. \quad (2.3)$$

Die **momentane Geschwindigkeit** \bar{v}_m wird zu einem festen Zeitpunkt auf einem Streckenabschnitt untersucht. Für N Messwerte und die momentane Geschwindigkeit eines Fahrzeugs v_{mi} ergibt sich

$$\bar{v}_m = \frac{1}{N} \sum_{i=1}^N v_{mi}. \quad (2.4)$$

Die **Zeitlücke** beschreibt den zeitlichen Abstand zweier aufeinander folgender Fahrzeuge beim Durchfahren eines Messquerschnitts. Meist wird die Zeitlücke von Bug zu Bug als Bruttozeitlücke gemessen. Die Nettozeitlücke ist die Lücke vom Heck des vorausfahrenden zum Bug des nachfolgenden Fahrzeugs.

Der **räumliche Abstand** bezeichnet analog den Weg zwischen zwei Fahrzeugen. Wieder wird in Brutto- und Nettoweglücke unterschieden.

Eine weitere Kenngröße ist die **Belegung**, die definiert wird als der prozentuale Anteil der Zeit, während der ein Detektor von einem Fahrzeug belegt ist. Sie hängt davon ab, wie lang die Fahrzeuge und wie groß die Lücken zwischen ihnen sind [49].

2.1.2 Lichtsignalanlagen

Eine **Lichtsignalanlage (LSA)** heißt im Verkehrsrecht und der Straßenverkehrsordnung *Lichtzeichenanlage* und im umgangssprachlichen Gebrauch meist *Ampel*. Zu einer Lichtsignalanlage gehören als Signalgeber drei farbige Leuchten sowie Lautsprecher und Vibrationsflächen für Blindensignale. Neben den Signalgebern sind auch Maste, Verkabelung und Steuergerät Teil der LSA. [13], [43]

Die *Richtlinien für Lichtsignalanlagen* - kurz RiLSA - sind das in Deutschland gültige technische Regelwerk, das für Entwurf, Ausführung und Betrieb von Lichtsignalanlagen Regeln und Empfehlungen aufstellt. RiLSA wird von der Forschungsgesellschaft für Straßen- und Verkehrswesen herausgegeben. Die aktuelle Version ist Ausgabe 2015. Anhand von RiLSA [13] wird nun die Funktion von Lichtsignalanlagen erklärt.

Der Zweck von Lichtsignalanlagen ist es, die Verkehrssicherheit und die Qualität des Ver-

kehrsablaufs zu verbessern. Sie kontrollieren die Verkehrsabwicklung an Knotenpunkten, Streckenabschnitten und in Verkehrsnetzen.

Ein Signalprogramm ist eine Festlegung der Signalfolgen hinsichtlich Dauer und Zuordnung. Dabei ist die für Kraftfahrzeuge vorgegebene Signalfolge *GRÜN-GELB-ROT-ROT/GELB-GRÜN*. Während einer Phase ändert sich der Zustand der Signalisierung nicht. Für die Phaseneinteilung sind zunächst Verkehrsströme so zusammenzufassen, dass verträgliche und bedingt verträgliche Verkehrsströme identifiziert werden. Verkehrsströme, die verträglich sind, haben keine gemeinsamen Konfliktflächen. Bedingt verträgliche Verkehrsströme können Konfliktflächen mit dem Gegenverkehr oder Fußgängern haben. Um Freigabezeiten des einen Stroms von denen des folgenden Stroms sicher zu trennen, sollen Zwischenzeiten eingehalten werden, die eine sichere Kreuzungsräumung ermöglichen.

Der Ablauf eines Signalprogramms ist durch ein Steuerungsverfahren festgelegt. Solche Steuerungsverfahren können unterschiedlich beschaffen sein. Grundsätzlich werden sie in verkehrsabhängige Verfahren und durch Festzeit gegebene Verfahren unterteilt. Dabei ist eine Festzeitsteuerung unabhängig von der Verkehrsnachfrage und fest mit Phasendauer und -reihenfolge vorgegeben. Verkehrsabhängige Steuerungsverfahren können makroskopisch oder mikroskopisch sein. Während makroskopische Verfahren anhand makroskopischer Verkehrskenngrößen wie mittlerer Staulänge, Verkehrsdichte oder Emissionsgrenzwerten definierbar sind, sind mikroskopische Steuerverfahren zumeist in ein makroskopisches Pendant eingebettet und in dessen Rahmen für Veränderungen verantwortlich. So ist es zum Beispiel möglich, die Phasen im vorgegebenen Plan untereinander auszutauschen oder die Freigabezeiten je nach Bedarf zu verändern. In Abbildung 2.1 ist ein Beispiel für einen Phasenfolgeplan zu sehen.

Dort sind drei Phasen mit den zugehörigen Verkehrsströmen dargestellt. Pfeile zeigen die möglichen Phasenübergänge. Es ist zu sehen, dass die Phasenfolge nicht fest vorgegeben ist. Es ist möglich von Phase 1 direkt nach Phase 2 oder Phase 3 zu wechseln. Von beiden ist wiederum der Wechsel zu Phase 1 möglich. Der Übergang von Phase 3 zu Phase 2 hingegen ist nur in diese eine Richtung möglich. Ist die Lichtsignalanlage in Phase 2, erfolgt zwangsläufig der Wechsel zu Phase 1.

Oft werden tageszeit- und wochentagsabhängig unterschiedliche Signalprogramme verwendet. Wird das Signalprogramm gewechselt, ist zusätzlich ein klar definiertes Umschaltverfahren erforderlich. Zudem sollten benachbarte Gebiete nach Möglichkeit an das verwendete Signalprogramm gekoppelt werden, um systematische Wartezeiten zu vermeiden. Die Steuerung mittels einer *grünen Welle* basiert beispielsweise auf dieser Überlegung. Weitere Aspekte, die je nach Einsatzort der Lichtsignalanlage ins Gewicht fallen, sind die Einbindung von Fußgängern, das Regeln des Radverkehrs sowie die Bevorzugung des ÖPNV. [13]

In dieser Arbeit werden Szenarien betrachtet, die eine einzelne Kreuzung mit Lichtsignalanlage vorsehen. Die Lichtsignalanlage ist dabei mit einer Festzeitsteuerung ausgestattet.

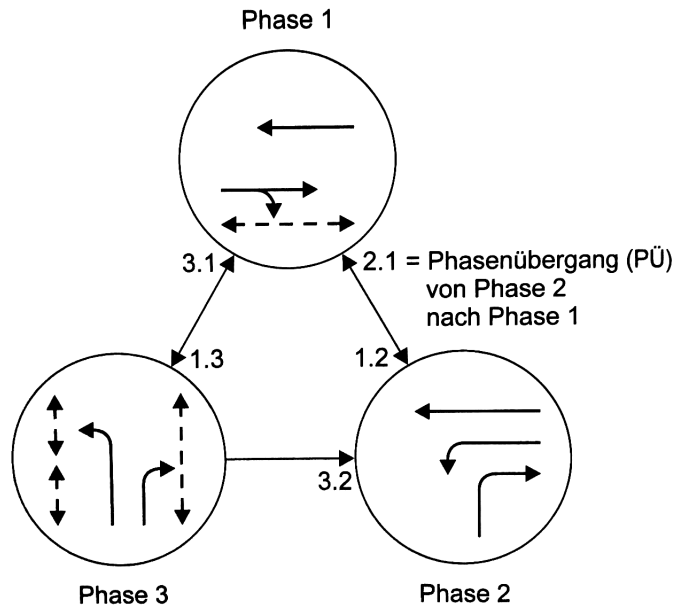


Abbildung 2.1: Beispiel für einen Phasenfolgeplan mit drei Phasen, aus [13].

2.2 Kommunikation und Sensoren

In diesem Abschnitt wird die Funktionsweise der Car-2-X-Kommunikation und des verwendeten Bluetooth-Verfahrens beschrieben. Dafür werden deren Funktionalitäten erörtert und dargestellt, was es für die verwendeten Verfahren zu beachten gilt.

2.2.1 Car-2-X-Kommunikation

Die *Car-2-X-Kommunikation* (C2X) beschreibt die Kommunikation zwischen Fahrzeugen untereinander (*Car-2-Car*) und die zwischen Fahrzeug und Infrastruktur (*Car-2-Infrastructure*). Ferner können auch die Kommunikation zwischen Fahrzeug und Fußgänger sowie die zwischen Fahrzeug und Netzwerken unter den Begriff gefasst werden. Die C2X-Kommunikation soll den Rahmen dafür liefern, kooperative Verkehrsdaten zwischen Verkehrsteilnehmern und Infrastruktur auszutauschen. Sie bietet neue Chancen für automatisiertes Fahren, Verkehrssicherheit und -effizienz. [7]

So wirbt Mercedes Benz etwa damit, dass C2X-Kommunikation es ermöglicht, Gefahren zu registrieren, bevor diese sichtbar sind. Mittels C2X können Fahrzeuge in der Umgebung vor erfassten Gefahren gewarnt werden. Auch zur Verkehrseffizienz kann C2X-Kommunikation beitragen. Werden Unfälle vermieden, so bleiben auch durch sie bedingte Verkehrsbehinderungen aus. So kann der Verkehr freier und effizienter fließen. [33]

Das *CAR 2 CAR Communication Consortium (C2C-CC)* ist eine industriegetriebene Nonprofitorganisation aus Fahrzeugherstellern, Zulieferern, Forschungsorganisationen und weiteren Partnern, die sich als Ziel gesetzt hat, kooperative intelligente Transportsysteme voranzutreiben und einen europäischen und internationalen Standard für C2X zu schaffen. [2]

Den geforderten Standards folgend basiert C2X-Kommunikation auf WLAN-Kommunikation. Die beiden Hauptpfeiler des Kommunikationssystems sind einerseits die räumliche und zeitliche Verbreitung von Informationen zwischen Fahrzeugen und andererseits die Bereitstellung eines Systems zur Nachrichtenübertragung durch Paketvermittlung. Zur Kommunikation wird der WLAN-Standard IEEE 802.11p genutzt. Fahrzeuge sind mit einer kommunikationsfähigen On-Board Unit ausgestattet, während Infrastrukturkomponenten über Roadside Units zum Nachrichtenaustausch verfügen.

Über Nachrichten können Sensordaten der Verkehrsteilnehmer ausgetauscht werden. Wichtige Parameter sind dabei etwa Positionsdaten, Geschwindigkeit und Fahrtrichtung. Es gilt hier feste Datenstrukturen für bestimmte Nachrichtentypen zu definieren. [2], [3]

Aktuell ist der Anteil der Fahrzeuge mit C2X-Ausstattung noch gering. Da die neue Technologie sich noch verbreiten und durchsetzen muss, gibt es nur Prognosen darüber, wie sich die Ausstattung in Zukunft verändert. So schätzt Daimler im Projekt *DRIVE C2X* im Jahr 2014 die Ausstattungsgrade für die Jahre 2020 und 2030 mit je einer optimistischen und einer pessimistischen Ausstattungsrate wie folgt: [42]

	2020	2030
pessimistisch	3,75%	26,29%
optimistisch	15,32%	100%

Tabelle 2.1: Schätzung des C2X-Ausstattungsgrades im Jahr 2020 und 2030, *DRIVE C2X*, Daimler, 2014.

ABI research [1] stellte 2013 die Schätzung auf, dass 2027 70% C2X-fähige Fahrzeuge auf den Straßen sein werden, nachdem in 2018 über 10% Penetrationsrate erreicht werden. Da die ersten europäischen C2X-fähigen Fahrzeuge erst seit April 2016 auf den Straßen sind, erscheinen diese Prognosen fraglich [21].

Die Anwendung von auf C2X-Technologie basierender Verkehrssteuerung hängt also maßgeblich davon ab, wie sich die Technologie in den nächsten Jahren tatsächlich verbreitet. Es ist daher wichtig, die Abhängigkeit solcher Modelle von den Ausstattungsgraden zu betrachten. Wenn die Ausstattungsgrade zukünftig tatsächlich steigen, sind bereits Strategien entwickelt, wie der neuartige Verkehr gesteuert werden kann.

Für die C2X-Kommunikation sind Nachrichtentypen definiert. So sind *Cooperative Awareness Messages (CAM)* dafür zuständig, Empfänger in der Umgebung über die Präsenz des sendenden Fahrzeugs zu informieren. Diese Nachrichten zeichnen sich durch ihre generische Struktur aus und werden genutzt, um den Status aller umliegenden

C2X-Kommunikationsteilnehmer zu erfassen. Neben den CAM-Nachrichten gibt es auch Nachrichtentypen speziell für Umgebungsmeldungen oder Informationen über die Lichtsignalanlage.

Um mit C2X Informationen zur Geschwindigkeit zu übermitteln, kann in wiederholten CAM-Nachrichten die Fahrzeugposition und die daraus resultierende Geschwindigkeit versendet werden. Die Fahrzeugposition wird dafür zunächst mittels GPS (Global Positioning System) bestimmt. Das satellitenbasierte Verfahren liefert auf 3-5 m genaue Ergebnisse. [16], [18]

Verkehrssteuerung mit C2X-Kommunikation wird zum Beispiel im autonomen Kreuzungsmanagement genutzt. Dabei nimmt jedes ankommende Fahrzeug mit einer Roadside Unit am Knotenpunkt Kontakt auf. Diese beinhaltet einen Server, der die Anfragen der Fahrzeuge bearbeitet. Das Fahrzeug fordert dann die Durchfahrtserlaubnis mit der geplanten Fahrtrichtung an. Der Server simuliert die Trajektorien, um mögliche Kollisionen zu vermeiden. Ist eine sichere Durchfahrt möglich, erteilt der Server die Fahrerlaubnis. Andernfalls muss auf die Durchfahrt verschoben werden, bis sie kollisionsfrei möglich ist.

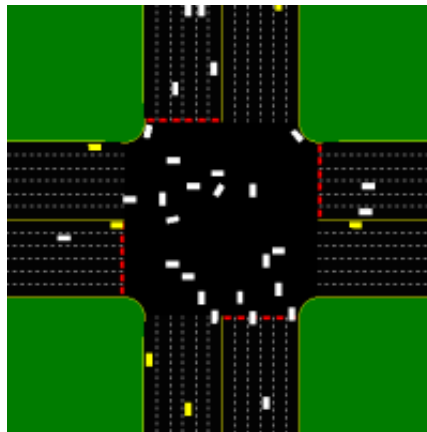


Abbildung 2.2: Beispiel für autonomes Kreuzungsmanagement. Simulation aus [10].

So können viele Fahrzeuge die Kreuzung ohne Stopp queren. In Abbildung 2.2 ist eine Simulation des Autonomen Kreuzungsmanagements zu sehen. Viele Fahrzeuge aus unterschiedlichen Richtungen queren hier gleichzeitig die Kreuzung.

Die eingesparten Abbrems- und Beschleunigungsvorgänge führen zu freierem Verkehrsfluss, weniger Kraftstoffverbrauch und reduzierten Emissionen. Diese Kreuzungssteuerung ist nur anwendbar, wenn alle Fahrzeuge mit C2X-Kommunikation ausgestattet sind. Andernfalls kann es serverseitig nicht erfasste Fahrzeuge im Kreuzungsbereich geben, die eine Aussage über sichere Durchfahrtszeitpunkte unmöglich machen. [10]

2.2.2 Bluetooth

Bluetooth entstand 1998 aus einer Studie, in der eine kostengünstige Möglichkeit für kabellosen und energiesparenden Datenaustausch gesucht wurde. Es handelt sich um einen offenen Industriestandard, der kostenlos und frei verfügbar ist. Dadurch kann eine weite Verbreitung der Technologie erreicht werden. Mit 802.15 wurde ein IEEE-Standard für Bluetooth geschaffen. Bluetooth ermöglicht die spontane drahtlose Vernetzung unterschiedlicher Geräte verschiedener Hersteller. Diese Eigenschaft kann für die Verkehrszustandsbestimmung ausgenutzt werden. [22]

Die Bluetooth-Penetrationsrate in Fahrzeugen ist dank moderner Freisprecheinrichtungen, Smartphones und Navigationssysteme steigend. Je nach Straßen- und Fahrzeugtyp beträgt sie zwischen 3% und 50%. [37]

Dabei unterscheiden sich die Bluetooth-Geräte anhand ihrer Reichweiten in drei Klassen: Klasse 3 hat eine Reichweite von bis zu 10 m, Klasse 2 bis 30 m und Klasse 1 mit einer großen Reichweite von bis zu 100 m.

Jedes Bluetooth-Gerät hat eine eindeutige ID, die 48 Bit lange MAC-Adresse. Im *Inquiry*-Zustand sucht der Sender nach Bluetooth-Empfängern, zu denen eine Verbindung aufgebaut werden kann. Erst nach erfolgreichem Aufbau einer Verbindung können Daten übertragen werden.

Hat der Sender einen Empfänger detektiert, werden IDs und Zeitstempel ausgetauscht. Auch wenn eine weitere Datenübertragung ausbleibt, können so Bluetooth-Geräte eindeutig identifiziert und ihre Detektionen zeitlich festgehalten werden. Wird ein Bluetooth-Empfänger als Infrastruktur der Straße angebracht, kann er vorbeifahrende Fahrzeuge mit Bluetooth-Geräten detektieren. Die Anzahl der Bluetooth-Detektionen hängt von der Geschwindigkeit des Fahrzeugs ab und wiederholt sich mit dem Inquiry-Prozess.

Der Inquiry-Vorgang hat eine festgelegte Dauer und wiederholt sich periodisch. Während zunächst 10,24 s empfohlen waren, konnten experimentell gute Ergebnisse auch für kürzere Inquiry-Scans von bis zu 1,28 s erzielt werden [38]. Im Folgenden soll orientiert an [24] eine übliche Inquiry-Zeit von 2,56 s angenommen werden.

Ist der Detektionsbereich r_{BT} des Senders bekannt, so kann die Geschwindigkeit des Empfängers geschätzt werden. Die Differenz aus letztem Detektionszeitpunkt t_{Last} und erstem Detektionszeitpunkt t_{First} des jeweiligen Empfängerfahrzeugs ist eine Schätzung der Durchfahrtszeit durch den Bluetooth-Detektionsbereich. Damit lässt sich die Geschwindigkeit grob schätzen als: [24], [40]

$$v_{BT} \approx \frac{r_{BT}}{t_{Last} - t_{First}}. \quad (2.5)$$

In Abbildung 2.3 ist dargestellt, wie die Bluetooth-Schätzung aus 2.5 anschaulich abläuft.

Dabei ist zu sehen, wie das Fahrzeug in den Detektionsbereich einfährt und die erste Detektion kurz danach stattfindet. Die Kreuze zeigen die Position des Fahrzeugs bei der Bluetooth-Detektion. Die letzte Detektion ist räumlich und zeitlich kurz vor dem

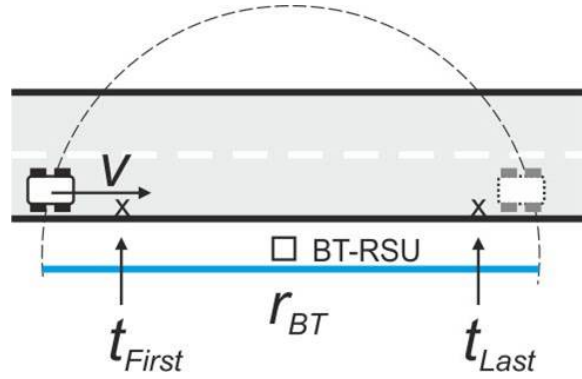


Abbildung 2.3: Ablauf der Bluetooth-Detektion mit erster und letzter Detektion eines Fahrzeugs. t_{First} : erste Detektion, t_{Last} : letzte Detektion, r_{BT} : Reichweite des Bluetooth-Empfängers, v : tatsächliche Geschwindigkeit, BT-RSU: Bluetooth Roadside Unit, aus [24].

Verlassen des Detektionsbereichs r_{BT} . Zusätzliche Detektionen zwischen diesen Punkten sind ebenfalls möglich.

Es lassen sich zwei generelle Eigenschaften der Geschwindigkeitsschätzung durch Bluetooth-Detektion festhalten. Da die erste Detektion stets nach Einfahrt in den Bluetooth-Detektionsbereich erfolgt und die letzte Detektion noch vor Ausfahrt aus dem Bereich stattfindet, wird die Durchfahrtszeit systematisch zu niedrig geschätzt. Daher wird die Geschwindigkeit bei festgelegter Reichweite des Empfängers r_{BT} stets überschätzt.

Zweitens hat die Schätzung für eine gegebene Reichweite ein Maximum, das durch die 2,56s-Periodizität des Inquiry-Vorgangs festgelegt ist. Weniger als diese Zeitspanne können zwei Detektionszeiten nicht voneinander abweichen. Für einen 30m großen Detektionsbereich ist die maximale detektierbare Geschwindigkeit

$$v_{BT,max} = \frac{30 \text{ m}}{2,56 \text{ s}} \approx 11,72 \text{ m/s} \approx 42,19 \text{ km/h.} \quad (2.6)$$

Für die weiteren Untersuchungen soll noch der Fall betrachtet werden, dass ein Fahrzeug detektiert wird und anschließend den Detektionsbereich verlässt, bevor eine zweite Detektion stattfindet. Besonders bei schnellen Fahrzeugen tritt dies auf, da die Wahrscheinlichkeit einer erneuten Detektion mit der Verweildauer im Detektionsbereich steigt. Um diese Fahrzeuge dennoch zu erfassen, soll ihre Geschwindigkeit pauschal mit der innerstädtischen Höchstgeschwindigkeit von $50 \text{ km/h} \approx 13,89 \text{ m/s}$ geschätzt werden.

Schließlich ist festzuhalten, dass diese Schätzung bei praktischer Anwendung Probleme birgt. So lässt sich der Bluetooth-Detektionsbereich nicht exakt festlegen, sondern ist als Maximalwert zu verstehen, der nur unter optimalen Bedingungen erreicht wird. Die

tatsächliche Reichweite liegt häufig unter dem angegebenen Wert. Neben Sendeleistung und Empfindlichkeit des Empfängers beeinträchtigen auch Hindernisse die Reichweite. Mauern und Metallteile sowie größere Objekte führen zu einer verminderten Leistung. [8],[37]

2.3 Verkehrssimulator SUMO

In diesem Abschnitt wird der für experimentelle Untersuchungen des Modells verwendete Verkehrssimulator SUMO vorgestellt. Die vorbereiteten Szenarien zur späteren simulativen Untersuchung werden hier erklärt. Auch das in SUMO intern verwendete Fahrzeugfolgemodell von Krauß wird beschrieben. Schließlich wird das Interface TraCI, mit dem SUMO in Echtzeit gesteuert werden kann, in seiner Funktionsweise dargestellt.

2.3.1 SUMO

SUMO - kurz für *Simulation of Urban MObility* - ist ein Open-Source-Verkehrssimulationspaket, das aus einer Simulationsanwendung und unterstützenden Werkzeugen besteht. Der Simulator findet Verwendung, um vielfältige Forschungsthemen im Verkehrsmanagement und in der Fahrzeugkommunikation zu untersuchen. SUMO ist eine mikroskopische Plattform zur Verkehrsflusssimulation. In SUMO können unterschiedliche Verkehrsteilnehmer wie Fußgänger, Radfahrer und Fahrzeuge dargestellt werden. Sie ist zeitdiskret und räumlich kontinuierlich. [29]

Seit 2002 die erste Version veröffentlicht wurde, hat sich der Umfang des Pakets beständig vergrößert. Heute gehört der Import von Straßennetzen in verschiedenen Formaten ebenso dazu wie die Generierung von Verkehrsnachfragen. Auch die High Performance Simulation einer einzelnen Kreuzung sowie das Interface TraCI zur Fernsteuerung von SUMO sind Bestandteil des Pakets. Die Weiterentwicklung von SUMO wird vor allem vom Deutschen Zentrum für Luft- und Raumfahrt vorangetrieben.

In SUMO werden Straßennetze als Graphen dargestellt. Die Knoten stellen Kreuzungen dar, die durch gerichtete Kanten mit Straßen versehen werden. Jede Kante hat eine festgesetzte Anzahl an Fahrbahnen. Attribute wie Position und Form der Kreuzung, Vorfahrtsregeln, Lichtsignalanlagen sowie Fahrbahngeometrien definieren Straßen und Kreuzungen weiter. Die Generierung neuer Netze erfolgt durch Werkzeuge zum Erstellen oder Importieren vorhandener Netze.

Fahrzeuge werden auf rein mikroskopischer Ebene dargestellt. Jedes Fahrzeug wird explizit angegeben und ist mindestens mit einer eindeutigen Kennung, seiner Abfahrtszeit und der zu fahrenden Route versehen. Zusätzlich können Geschwindigkeit, genutzte Fahrbahnen und Fahrzeugtypen definiert werden, die physikalische Eigenschaften des Fahrzeugs beschreiben. Eine weitere Möglichkeit, Fahrzeuge in die Simulation zu integrieren, besteht darin, Verkehrsflüsse zu definieren. Diese definieren das wiederholte

Einsetzen von Fahrzeugen in einem gegebenen Zeitintervall. Dabei wird die definierte Anzahl Fahrzeuge standardmäßig gleichverteilt eingesetzt. [48]

Die Anwendung *sumo* führt die zeitdiskrete Simulation mit Schrittweiten zwischen 1ms und 1s durch.

Zur Fahrzeugbewegung wird ein Fahrzeugfolgemodell verwendet, das die Fahrgeschwindigkeit vorgibt. Ist kein vorausfahrendes Fahrzeug oder sonstiges Hindernis vorhanden, wird die vordefinierte Wunschgeschwindigkeit angestrebt und konstant gehalten. Als Fahrzeugfolgemodell wird eine Modifikation des Modells von Krauß verwendet. Es nutzt die Eigengeschwindigkeit, die Geschwindigkeit des vorausfahrenden Fahrzeugs und den Abstand zwischen den beiden Fahrzeugen [29]. Das Modell wird im folgenden Abschnitt erklärt.

2.3.2 Fahrzeugfolgemodell von Krauß

Basierend auf der Grundannahme, dass sich Fahrzeuge in der Regel kollisionsfrei bewegen, erstellte Krauß 1998 ein Verkehrsflussmodell. Hauptcharakteristika des Modells sind Parameter, die typische Beschleunigungs- und Verzögerungsfähigkeiten des Fahrzeugs beschreiben. Das Modell von Krauß hat den Anspruch, den Geltungsbereich bestehender Modelle zu vergrößern, der oft nur bestimmte Szenarien umfasst.

Krauß unterscheidet grundlegend zwei Fahrsituationen: Die Fahrt in freiem Verkehrsfluss und die Fahrt in Interaktion mit anderen Verkehrsteilnehmern. Im freien Verkehr ist die Geschwindigkeit nur durch die Maximalgeschwindigkeit begrenzt $v_{max} \geq v$, die als Wunschgeschwindigkeit des jeweiligen Fahrers verstanden werden kann. Ist Interaktion mit anderen Verkehrsteilnehmern erforderlich, so ist der Fahrer bestrebt, Kollisionen zu vermeiden, und hält die Geschwindigkeit stets unterhalb einer maximalen sicheren Geschwindigkeit v_{safe} . Die Definitionen und Herleitungen folgen. [30]

Abbremsung b und Beschleunigung a sind begrenzt durch Randparameter $a, b > 0$, sodass für die zeitabhängige Geschwindigkeit $v = v(t)$ gilt

$$-b \leq \frac{dv}{dt} \leq a. \quad (2.7)$$

Für diskrete Zeitschritte Δt ergibt sich damit die Beschränkung

$$v(t + \Delta t) \leq \min\{v_{max}, v(t) + a\Delta t, v_{safe}\}. \quad (2.8)$$

Da b die maximale Abbremsung angibt, gilt für alle $v(t + \Delta t)$, dass

$$v(t + \Delta t) \geq v(t) - b\Delta t. \quad (2.9)$$

Auch v_{safe} muss diese Bedingung berücksichtigen. Es gilt

$$v_{safe}(t + \Delta t) \geq v(t) - b\Delta t. \quad (2.10)$$

Für die Interaktion zwischen zwei Fahrzeugen an Position x_f und x_l , wobei x_f Nachfolger von x_l ist, wird die Nettoraumlücke g abhängig von der Fahrzeuglänge l betrachtet

$$g = x_l - x_f - l. \quad (2.11)$$

Die zeitliche Änderung der Lücke hängt von den Geschwindigkeiten der beiden Fahrzeuge ab und es gilt

$$\dot{g}(t) = v_l(t) - v_f(t). \quad (2.12)$$

Um Kollisionen zwischen den Fahrzeugen zu vermeiden, genügt es sicherzustellen, dass diese Lücke zu jedem Zeitpunkt groß genug bleibt. Sei τ_{des} die gewünschte Relaxationszeit. Diese ist ein Maß dafür, wie sich ein Fahrer der gewünschten Geschwindigkeit annähert.

Für die gewünschte räumliche Lücke g_{des} gilt die Ungleichung

$$\frac{dg}{dt} \geq \frac{g_{des} - g}{\tau_{des}}. \quad (2.13)$$

Dabei können g_{des} und τ_{des} Funktionen sein. τ_{des} besteht hier aus einer individuell an den Fahrer angepassten Komponente τ und einer von der maximalen Abbremsung b abhängigen Komponente τ_b . Damit ist $\tau_{des} = \tau + \tau_b$.

Für die diskrete Simulation ergibt sich mit (2.12)

$$v_f(t + \Delta t) \leq v_l(t) + \frac{g(t) - g_{des}(t)}{\tau_{des}(t)}. \quad (2.14)$$

Die Position der Fahrzeuge ergibt sich in jedem Zeitschritt aus

$$x(t + \Delta t) = x(t) + v(t + \Delta t)\Delta t. \quad (2.15)$$

Für $\Delta t \rightarrow 0$ und $g_{des} \geq 0$ garantiert die Gleichung Sicherheit und Freiheit von Kollisionen. Für ein festes $\Delta t > 0$ muss die Bedingung hingegen gesondert überprüft werden. Die Raumlücke zur Zeit $t + \Delta t$ beträgt

$$g(t + \Delta t) = g(t) + \Delta t(v_l(t + \Delta t) - v_f(t + \Delta t)). \quad (2.16)$$

Mit (2.14) und $\xi(t) = g(t) - v_l(t)\Delta t$ ergibt sich

$$\xi(t + \Delta t) \geq \xi(t)\left(1 - \frac{\Delta t}{\tau_b + \tau}\right) + \Delta t \frac{g_{des}(t) - v_l(t)\Delta t}{\tau_b + \tau}. \quad (2.17)$$

Sicherheit kann garantiert werden, wenn $g \geq 0$ gilt. Hierfür reicht es, wenn $\xi(0) \geq 0$ und für alle $t > 0$

$$\Delta t \leq \tau, \quad (2.18)$$

$$g_{des} \geq v_l \Delta t. \quad (2.19)$$

Diese beiden Bedingungen formalisieren die Forderungen, dass die Relaxationszeit größer sein muss als der Zeitschritt und dass der gewünschte räumliche Abstand größer ist als die bei gleicher Geschwindigkeit im Zeitschritt zurückgelegte Strecke.

Der Geschwindigkeit sei eine zufällige Störung $\eta > 0$ hinzugefügt, die Abweichungen von dem optimalen Fahrverhalten erlaubt. Das gesamte Modell ist somit gegeben durch

$$v_{safe}(t) = v_l(t) + \frac{g(t) - g_{des}(t)}{\tau_b + \tau}, \quad (2.20)$$

$$v_{des}(t) = \min\{v_{max}, v(t) + a(v)\Delta t, v_{safe}(t)\}, \quad (2.21)$$

$$v(t + \Delta t) = \max\{0, v_{des}(t) - \eta\}, \quad (2.22)$$

$$x(t + \Delta t) = x(t) + v\Delta t. \quad (2.23)$$

In einem weiteren Schritt lassen sich auch Fahrspurwechsel in das Modell aufnehmen [30]. Für die Zwecke dieser Arbeit ist diese Variante jedoch ausreichend.

Im Verkehrssimulator SUMO findet das Krauß-Modell Anwendung, da es sich durch Einfachheit und geringen Rechenaufwand auszeichnet [29].

Das Modell wurde für die Verwendung in SUMO in zweierlei Hinsicht abgeändert. In SUMO können Fahrzeuge in der gleichen Simulation mit unterschiedlichen Abbremsfähigkeiten ausgestattet werden. Für diesen Fall können im ursprünglichen Modell von Krauß Kollisionen entstehen. Daher muss sichergestellt werden, dass keine Verletzungen der Sicherheitsrichtlinien durchgeführt werden können.

Zweitens werden die Positionen der Fahrzeuge in SUMO berechnet, indem in jedem Zeitschritt die Fahrzeuggeschwindigkeiten als konstant angenommen werden. Im Originalmodell von Krauß hingegen wird die Beschleunigung als konstant angenommen. Daher muss die Formel zur Bestimmung der sicheren Geschwindigkeit angepasst werden, um weiterhin Sicherheit zu gewährleisten. Hierfür wurden einige der stetigen Terme diskretisiert. [48]

2.3.3 Szenarien

Ein SUMO-Szenario besteht aus einem Straßennetz und einer darauf zugeschnittenen Verkehrsnachfrage. Als Schnittstelle fungiert dabei eine .sumocfg-Konfigurationsdatei, die zum Starten der Anwendung benötigt wird und die die Bestandteile der Simulation definiert. Dort werden XML-Dateien für das zugrundeliegende Netz und die Definition

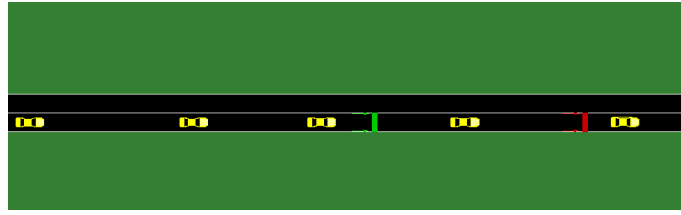


Abbildung 2.4: Minimales Szenario mit Bluetooth-Bereich von 30m.

der Verkehrsnachfrage eingebunden. Optional können zusätzliche Verkehrsteilnehmer, Detektoren und Output-Dateien als weitere Elemente eingefügt werden.

Zur simulativen Untersuchung der Problemstellung wurden mehrere Testszenarien vorbereitet. Diese sind so gewählt, dass in akademischen Szenarien bestimmte Fragestellungen untersucht werden können, aber dank komplexerer realitätsnaher Szenarien auch eine wirklichkeitsnahe Erprobung ermöglicht wird.

Das erste **Szenario Minimal** ist ein Minimalbeispiel, in dem auf einem einzelnen Straßenabschnitt ohne Kurven oder Kreuzungen elementare Eigenschaften des Verkehrs dargestellt werden können. In Abbildung 2.4 ist das Szenario zu sehen.

Dieses Szenario enthält keinen Knoten und ermöglicht so, elementare Beobachtungen über die Funktionalität des Modells zu machen. Dabei ist der Verkehrsfluss so erzeugt, dass die gefahrene Geschwindigkeit über einen festgesetzten Zeitraum konstant gehalten und dann stückweise verringert wird. Solch ein Verkehrsfluss ist wenig realistisch, ermöglicht aber, den Einfluss der Geschwindigkeit im Detail zu erforschen. Zusätzlich wird auch die Verkehrsstärke schrittweise erhöht. Dafür wird zunächst eine geringe, dann eine mittlere und zuletzt eine hohe Verkehrsstärke angenommen. In 100 s werden zunächst 10, dann 20 und schließlich 40 Fahrzeuge eingesetzt. Das entspricht 360, 720 bzw. 1440 Fahrzeugen pro Stunde. Nachdem langsame Fahrzeuge eingesetzt wurden, sind Pufferzeiten ohne Einsetzen weiterer Fahrzeuge vorgesehen. So kann sichergestellt werden, dass die langsamen Fahrzeuge die Simulation verlassen haben, bevor schnellere Fahrzeuge eingesetzt werden und es zu Rückstau käme.

Der Verkehrsfluss ist in Tabelle 2.2 dargestellt.

Die Geschwindigkeitsreduktion findet in Schritten von 2 m/s von 14 auf 2 m/s statt. Dabei entspricht die Geschwindigkeit der mittleren Wunschgeschwindigkeit. Zur realistischeren Simulation wird die Geschwindigkeit der einzelnen Fahrzeuge gestreut. Dazu wird die Normalverteilung mit Erwartungswert $\mu = 1$ und Standardabweichung $\sigma = 0,1$ verwendet. Sie wird bei 0,5 und 2 abgeschnitten, um extreme Abweichungen von der mittleren Geschwindigkeit zu vermeiden. Abschließend wird sie auf die jeweilige Wunschgeschwindigkeit skaliert.

Die schrittweise Verringerung der Geschwindigkeit wird für alle drei Fälle wiederholt, sodass Daten für alle möglichen Kombinationen von Geschwindigkeit und Verkehrsstärke

Zeit [s]	Anzahl Fahrzeuge	Geschwindigkeit [m/s]
0 - 100	10	14
100 - 200	10	12
⋮	⋮	⋮
600 - 700	10	2
1000 - 1100	20	14
1100 - 1200	20	12
⋮	⋮	⋮
1600 - 1700	20	2
2000 - 2100	40	14
2100 - 2200	40	12
⋮	⋮	⋮
2600 - 2700	40	2

Tabelle 2.2: Verkehrsfluss für Szenario Minimal. Schrittweise werden Geschwindigkeit und Fahrzeuganzahl verändert.

gesammelt werden können. Allgemeine Aussagen über die Abhängigkeit der Ergebnisse von Geschwindigkeit und Verkehrsstärke können hier getroffen werden. Zudem werden Grenzfälle abgedeckt, die in anderen Szenarien kaum auftreten. Dazu gehören beispielsweise besonders niedrige Geschwindigkeiten bei niedriger Verkehrsstärke.

Im nächsten **Szenario RiLSA** ist ein Knoten Gegenstand der Untersuchungen. Das Szenario ist dabei dem ersten Beispiel in [14], der Beispielsammlung zu den Richtlinien für Lichtsignalanlagen, nachempfunden. Es besteht aus einer einzelnen Kreuzung, die sich durch ihre simpel gehaltene Geometrie auszeichnet. Am Knotenpunkt treffen vier Kreuzungsarme aufeinander, die rechtwinklig in alle vier Himmelsrichtungen verlaufen und symmetrische Geometrien haben, wie in Abbildung 2.5 dargestellt.

An der Kreuzung steuert eine Lichtsignalanlage den Verkehrsfluss. Auch dieses Szenario ist als akademisches Beispiel einzuordnen. Es ermöglicht Betrachtungen auf einer überschaubaren Kreuzung, die keine richtungsabhängigen Eigenschaften aufweist. Geometrische Besonderheiten, wie sie in der Realität an jedem Knoten vorkommen können, bleiben zunächst außen vor. Der Verkehrsfluss ist hier annähernd konstant über den gesamten Simulationszeitraum gewählt, sodass kreuzungsspezifische Phänomene und Abhängigkeiten von der Lichtsignalanlage untersucht werden können und es nur zu geringen Schwankungen durch Geometrie und Wechsel der Verkehrsnachfrage kommt.

Die Lichtsignalanlage folgt dabei einer Festzeitsteuerung. Auf der Nord-Süd-Achse beträgt die GRÜN-Phase 12 s, auf der Ost-West-Achse beträgt sie 40 s. Nach der GRÜN-Phase folgt jeweils eine drei Sekunden andauernde GELB-Phase sowie eine ROT-Phase von 7 s. Nach insgesamt 72 s wiederholt sich der Phasendurchlauf. In Tabelle 2.3 ist die

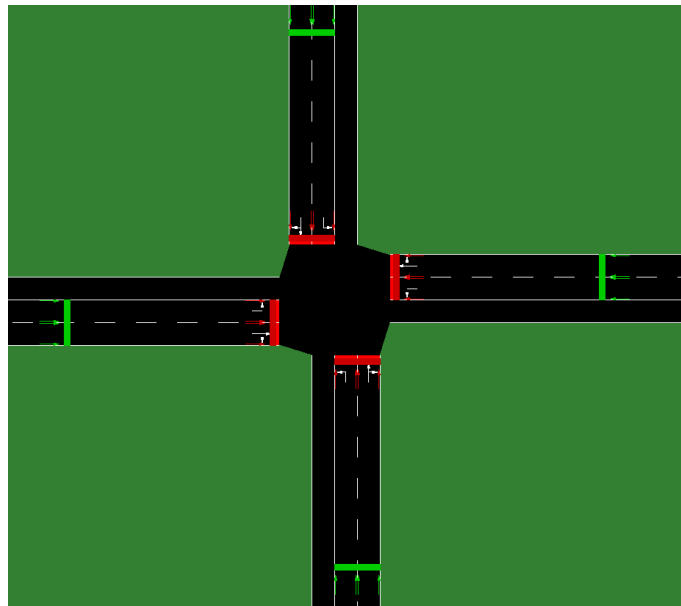


Abbildung 2.5: RiLSA Szenario mit vier Kreuzungsarmen.

Zeitdauer [s]	Phasennummer	Phase
7	0	
40	1	
3	2	
7	3	
12	4	
3	5	

Tabelle 2.3: Festzeitsteuerung für Szenario RiLSA.

Festzeitsteuerung dargestellt.

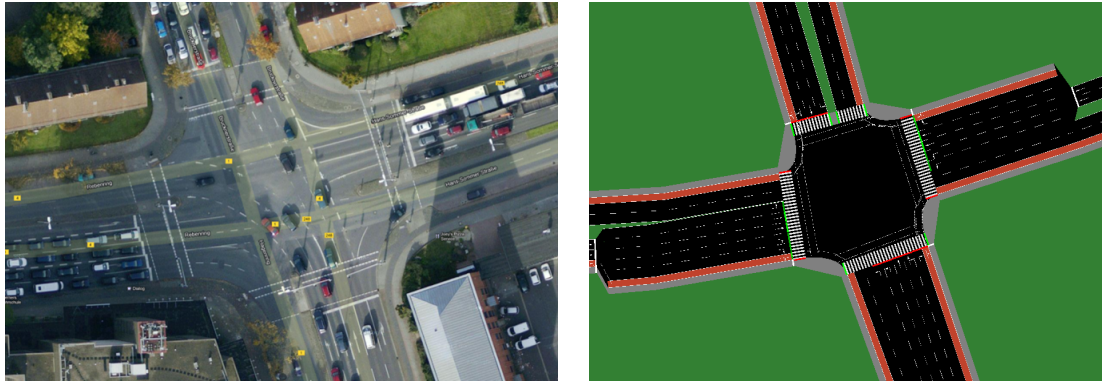


Abbildung 2.6: Szenario Forschungskreuzung in Braunschweig. Links: Reale Kreuzung, aus [50], rechts: Nachbildung als Netz in SUMO.

Die Verkehrsnachfrage ist dabei so gewählt, dass es eine stärker befahrene Ost-West-Achse und eine schwächer befahrene Nord-Süd-Achse gibt. Auf der stärker befahrenen Achse starten aus beiden Richtungen 1000 Fahrzeuge in der Simulationszeit von einer Stunde. Die Verkehrsnachfrage ist so gewählt, dass in einer Phase der Lichtsignalanlage der gesamte Rückstau gerade noch abgebaut werden kann. Da die Fahrzeuge probabilistisch verteilt werden, kann es hier zu Ausnahmen kommen, die sich aber schnell einpendeln. Auf der schwächer befahrenen Achse starten 300 Fahrzeuge aus jeder Richtung. Auch wenn die Lichtsignalphase hier kürzer ist, ist ihre Dauer in der Regel ausreichend, sodass alle Fahrzeuge die Kreuzung queren können.

Zur Modellierung der Verkehrsnachfrage werden Verkehrsflüsse mit Anfangs- und Endkante definiert. Da das Szenario nur aus je einer Kante in Nord-, Ost-, Süd- und Westrichtung besteht, genügt es, die Ausgangs- und Zielrichtung zu definieren. Für Fahrzeuge aus östlicher und westlicher Richtung werden je drei Verkehrsflüsse definiert. Dabei fahren je 600 Fahrzeuge geradeaus - also von Ost nach West und in entgegengesetzter Richtung - und je 200 Fahrzeuge biegen nach rechts und links ab. Für Fahrzeuge aus nördlicher und südlicher Richtung werden Verkehrsflüsse geradeaus, nach links und nach rechts von jeweils 100 Fahrzeugen definiert.

Das dritte **Szenario BS Forschungskreuzung** dient schließlich dazu, eine möglichst realistische Anwendung für das Modell zu schaffen. Dabei wird die Forschungskreuzung Hagenring / Rebenring / Hans-Sommer-Straße / Brucknerstraße in Braunschweig verwendet. Sie wird mit ihren realen Geometrien im Verkehrssimulator nachgebildet. In Abbildung 2.6 sind die reale Kreuzung und ihre Nachbildung in SUMO zu sehen.

Anders als im Szenario RiLSA sind die Kreuzungsarme hier nicht rechtwinklig und gerade, sondern haben unterschiedliche Geometrie. Längen, Winkel und Zahl der Fahr-

spuren unterscheiden sich dabei. Vorlage ist dabei stets die reale Kreuzung.

Aus früherer Forschungsarbeit existieren auch Videoaufzeichnungen des Verkehrsflusses an der Kreuzung. Daher kann auch die gemessene Verkehrsnachfrage hier realitätsgetreu abgebildet werden. Für die einzelnen Kreuzungsarme werden dabei 364 Fahrzeuge aus nördlicher, 586 aus östlicher, 645 aus südlicher und 472 aus westlicher Richtung eingesetzt. Die Simulationszeit beträgt dabei eine Stunde. Dabei ist jedoch im Gegensatz zu Szenario RiLSA nicht von einer konstanten Verkehrsstärke auszugehen. Die Fahrzeuge sind hier gemäß ihres tatsächlichen Eintreffens im Messbereich auch in der Simulation mit einem festen Einsetzzeitpunkt ausgestattet.

Die Lichtsignalanlage ist dabei wieder mit einer Festzeitsteuerung versehen, wie in Tabelle 2.4 gezeigt. Tatsächlich ist an der Kreuzung eine verkehrsabhängige Steuerung implementiert, die sich im Rahmen der hier verwendeten Festzeitsteuerung bewegt. Somit ist dies eine Annäherung an die tatsächlich aufgetretenen Phasen.

Dieses Szenario eignet sich, um die Anwendung im realitätsnahen Fall zu untersuchen.

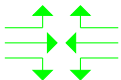
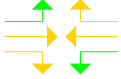
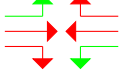

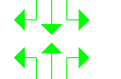
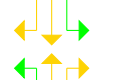


Zeitdauer [s]	Phasennummer	Phase
31	0	
4	1	
6	2	
4	3	
31	4	
4	5	
6	6	
4	7	

Tabelle 2.4: Festzeitsteuerung für Szenario BS Forschungskreuzung.

2.3.4 TraCI

TraCI - kurz für *Traffic Control Interface* - ist ein Interface zur Kopplung von Verkehrs- und Netzwerksimulatoren. Insbesondere ist es mit SUMO kompatibel. Es wurde 2006 von Axel Wegener et al. an der Universität Lübeck entwickelt mit dem primären Ziel, Fahrzeug-Ad-Hoc-Netzwerke zu modellieren. Solche Netzwerke ermöglichen die Kommunikation zwischen Fahrzeugen und Infrastruktur, wie in Abschnitt 2.2.1 vorgestellt. Hierfür sollen Verkehrs- und Netzwerksimulationen derart verknüpft werden, dass sich das Fahrverhalten zur Simulationszeit verändern lässt. Im Gegensatz zu vordefinierten Trajektorien, die vor Simulationsstart festgelegt werden, ist es mit TraCI möglich, die Fahrzeuge während der Simulation zu steuern und deren Trajektorien zur Laufzeit zu verändern. [29], [46]

Ein solches Verhalten kann in Fahrzeug-Ad-Hoc-Netzwerken durch Nachrichten getrig-

gert werden. Geht beispielsweise eine Unfallmeldung ein, kann die Route des Fahrzeugs neu berechnet werden, um die Unfallstelle zu vermeiden. Ist dies nicht möglich, kann die Geschwindigkeit nahe der Unfallstelle aus Sicherheitsgründen frühzeitig reduziert werden.

Um solche Situationen zu modellieren, bietet TraCI die Möglichkeit, in jedem Zeitschritt verschiedene Daten wie Fahrzeugpositionen und Geschwindigkeiten aus der Simulation zu übertragen. Diese Informationen sind auch außerhalb der Untersuchung von Fahrzeug-Ad-Hoc-Netzwerken nützlich, wann immer Fahrzeugdaten zur Laufzeit ausgelesen oder verändert werden sollen. [29], [46]

In der vorliegenden Arbeit wird TraCI genutzt, um die Simulation in SUMO zu kontrollieren. Dabei werden Detektoren ausgelesen und Fahrzeugdaten sowie der Zustand der Lichtsignalanlage abgefragt. Die mit TraCI versendeten Nachrichten forcieren den nächsten Simulationsschritt oder fordern Informationen aus der Simulation an.

3 Bayes'sche Wahrscheinlichkeit

In diesem Kapitel wird die theoretische Grundlage für das Konzept Bayes'scher Wahrscheinlichkeit im Allgemeinen und Bayes'scher Netze im Speziellen gelegt. Zunächst wird erklärt, wie sich klassische und Bayes'sche Wahrscheinlichkeitstheorie unterscheiden und was sie motiviert. Dabei wird eine kurze Zusammenfassung der im Folgenden genutzten wahrscheinlichkeitstheoretischen Zusammenhänge gegeben. Dann werden Bayes'sche Netze als grafische Repräsentationen der probabilistischen Beziehungen zwischen Variablen vorgestellt und erläutert, wie sie für Inferenzbetrachtungen angewendet werden können. Ein weiterer Abschnitt befasst sich mit Datenfusion und den damit verfolgten Zielen. Der letzte Abschnitt ist der Modellierung einer Verkehrszustandsschätzung mit Bayes'schen Netzen aus [24] gewidmet, auf der in dieser Arbeit aufgebaut wird. Das Modell wird mit besonderem Fokus auf die Modellierung des Bayes'schen Netzes zur Datenfusion vorgestellt. Schließlich wird erklärt, wie die Wahrscheinlichkeitsverteilungen bestimmt und mittels Schätzern ausgewertet werden können. Im nachfolgenden Kapitel werden diese theoretischen Betrachtungen aufgegriffen und für die praktische Modellierung eines Bayes'schen Netzes genutzt.

3.1 Bayes'sche und klassische Wahrscheinlichkeitstheorie

Bayes'scher und klassischer Wahrscheinlichkeitstheorie liegt eine unterschiedliche Definition des Wahrscheinlichkeitsbegriffs zugrunde. Daher werden zunächst verschiedene philosophische Ansätze betrachtet, die das Konzept der Wahrscheinlichkeit beschreiben. Neapolitan stellt in [35] drei Definitionen vor:

Die mengentheoretische Definition der Wahrscheinlichkeit geht auf A. N. Kolmogorov im Jahr 1933 zurück. Er axiomatisiert den Wahrscheinlichkeitsbegriff durch die Einführung einer Wahrscheinlichkeitsfunktion, die den Ereignissen in dem Ergebnisraum eines Zufallsexperiments Wahrscheinlichkeiten zuordnet.

Definition 1 (Diskrete Wahrscheinlichkeitsfunktion und Wahrscheinlichkeitsraum).

Gegeben sei ein Ergebnisraum $\Omega = \{e_1, e_2, \dots, e_n\}$ mit n verschiedenen Elementen. Sei $\mathcal{P}(\Omega)$ die zugehörige Potenzmenge, das heißt die Menge, die alle Teilmengen von Ω enthält.

Die Wahrscheinlichkeitsfunktion P ordnet jedem Ereignis $E \subseteq \mathcal{P}(\Omega)$ einen reellen Wert $P(E)$ zu und erfüllt dabei die folgenden Bedingungen:

1. $0 \leq P(\{e_i\}) \leq 1$ für $1 \leq i \leq n$.
2. $P(\{e_1\}) + P(\{e_2\}) + \dots + P(\{e_n\}) = 1$.
3. Für jedes Ereignis $E = \{e_{i_1}, e_{i_2}, \dots, e_{i_k}\}$ gilt

$$P(E) = P(\{e_{i_1}\}) + P(\{e_{i_2}\}) + \dots + P(\{e_{i_k}\}).$$

Das Tupel (Ω, P) wird Wahrscheinlichkeitsraum genannt.

Das von J. M. Keynes 1921 entwickelte Indifferenzprinzip schlägt weiter vor, Ereignisse als gleich wahrscheinlich anzunehmen, wenn kein Grund vorliegt, das eine oder andere Ereignis vorzuziehen. n solchen Ereignissen wird jeweils die Wahrscheinlichkeit $1/n$ zugeordnet. Daher wird diese Wahrscheinlichkeitsdefinition auch als Definition durch das Verhältnis bezeichnet.

Ein Elementarereignis ist ein Ereignis, das aus genau einem Element besteht. Für Elementarereignisse, die unterschiedlich wahrscheinlich sind - etwa auf welcher Seite eine Reißzwecke landet, nachdem sie in die Luft geworfen wurde - kann mit dieser Wahrscheinlichkeitsdefinition jedoch keine Aussage getroffen werden.

Hierfür kann die Wahrscheinlichkeit in einer zweiten Definition als relative Häufigkeit definiert werden. Konträr zum Indifferenzprinzip werden hierbei Wahrscheinlichkeiten nicht für jedes einzelne Experiment, sondern über die gesamte Versuchsreihe gemittelt. Für den Wurf der Reißzwecke bedeutet dies: Wird das Experiment n mal wiederholt und liegt k mal Kopf oben, so kann die Wahrscheinlichkeit mit $P(\{\text{Kopf}\}) \approx k/n$ angenähert werden. Sei n_K die Anzahl der Experimente mit Ergebnis Kopf. Die exakte Wahrscheinlichkeit wird mit dem Grenzwert

$$P(\{\text{Kopf}\}) = \lim_{m \rightarrow \infty} \frac{n_K}{m} \quad (3.1)$$

erreicht.

Dieses Prinzip erreicht seine Grenzen, wenn Experimente nicht beliebig oft durchführbar sind. Soll etwa eine Wahrscheinlichkeit für Gewinn und Niederlage bei einem Sportereignis angegeben werden, so ist dieses unter genau den gleichen Bedingungen nicht wiederholbar. Dennoch erscheint dem subjektiven Empfinden nach ein bestimmter Ausgang oft *wahrscheinlicher* als andere. Um solche gefühlten Wahrscheinlichkeiten abzubilden, schlägt Neapolitan in [34] vor, Wahrscheinlichkeit als Maß des Glaubens an eine bestimmte Hypothese zu definieren. Diese Wahrscheinlichkeit, so Neapolitan weiter, stimmt für die meisten Personen mit der Wahrscheinlichkeitsdefinition der relativen

Häufigkeit und des Verhältnisses überein.

Die Bayes'sche Wahrscheinlichkeitstheorie folgt der Definition der Wahrscheinlichkeit als Maß des Glaubens an eine Hypothese. Dadurch ist sie in der Lage, auch unbekannten Parametern Wahrscheinlichkeitsverteilungen zuzuordnen. Die klassische Wahrscheinlichkeitstheorie nimmt Wahrscheinlichkeiten hingegen als physikalische Eigenschaften an. Hier wird Wahrscheinlichkeit zur unumstößlichen Tatsache, die oft nur durch wiederholte Experimente messbar gemacht werden kann. Der Bayes'sche Ansatz erlaubt hingegen intuitive Festlegungen der Wahrscheinlichkeit. [20], [28]

Ist der Begriff der Wahrscheinlichkeit hinreichend geklärt, so lassen sich damit bedingte Wahrscheinlichkeit und bedingte Unabhängigkeit definieren. Die Darstellung der Definitionen und Theoreme folgt dabei [35].

Definition 2 (Bedingte Wahrscheinlichkeit).

Seien E und F Ereignisse, sodass $P(F) \neq 0$. Dann ist die bedingte Wahrscheinlichkeit von E unter F gegeben als

$$P(E|F) = \frac{P(E \cap F)}{P(F)}.$$

Definition 3 (Bedingte Unabhängigkeit).

Zwei Ereignisse E und F sind bedingt unabhängig unter G , wenn $P(G) \neq 0$ und eine der folgenden Bedingungen erfüllt ist:

1. $P(E|F \cap G) = P(E|G)$ und $P(E|G) \neq 0$, $P(F|G) \neq 0$.
2. $P(E|G) = 0$ oder $P(F|G) \neq 0$.

Mit diesen Definitionen kann der Satz von Bayes formuliert werden, der das Kernstück der Bayes'schen Wahrscheinlichkeitstheorie bildet. Er ermöglicht es, Wahrscheinlichkeitsverteilungen auch für unbekannte Parameter anzugeben. Wahrscheinlichkeit kann damit für Aussagen und nicht nur für Experimente angegeben werden. [28]

Der Satz von Bayes ist nach seinem Entdecker Sir Thomas Bayes benannt und lautet nach [5] wie folgt:

Satz 1 (Satz von Bayes).

Sei E_1, E_2, \dots, E_n eine Zerlegung der Ereignismenge E und F eine beliebige Menge. Dann gilt für alle $i = 1, \dots, n$, dass

$$P(E_i|F) = \frac{P(F|E_i)P(E_i)}{\sum_{j=1}^n P(F|E_j)P(E_j)}. \quad (3.2)$$

Der Beweis folgt [35].

Beweis (Satz von Bayes).

Für $n = 1$ gilt nach der Definition der bedingten Wahrscheinlichkeit

$$P(E_1|F) = \frac{P(E_1 \cap F)}{P(F)} \text{ und } P(F|E_1) = \frac{P(F \cap E_1)}{P(E_1)}.$$

Daher ist nach Multiplikation

$$P(E_1|F)P(F) = P(F|E_1)P(E_1)$$

und damit

$$P(E_1|F) = \frac{P(F|E_1)P(E_1)}{P(F)}.$$

Für $n > 1$ ist $E = E_1 \cup E_2 \cup \dots \cup E_n$ eine disjunkte Vereinigung. Daher gilt

$$P(F) = \sum_{i=1}^n P(F|E_i)P(E_i).$$

Es folgt direkt

$$P(E_i|F) = \frac{P(F|E_i)P(E_i)}{P(F)} = \frac{P(F|E_i)P(E_i)}{\sum_{j=1}^n P(F|E_j)P(E_j)}.$$

□

Es seien noch Zufallsvariablen und Wahrscheinlichkeitsverteilungen definiert. Sie bilden eine wichtige Grundlage für die Theorie der Bayes'schen Netze.

Definition 4 (Zufallsvariable, Wahrscheinlichkeitsverteilung).

Sei (Ω, P) ein Wahrscheinlichkeitsraum.

1. Dann ist eine Zufallsvariable X eine Funktion auf Ω .
2. Die Wahrscheinlichkeitsfunktion P definiert die Wahrscheinlichkeitsverteilung $P(X = x)$.
3. Für zwei Zufallsvariablen X und Y ist $P(X = x, Y = y)$ die gemeinsame Wahrscheinlichkeitsverteilung von X und Y .

In der Praxis kann der Wahrscheinlichkeitsraum oft nicht direkt angegeben werden. Stattdessen werden Zufallsvariablen definiert und ihre probabilistischen Beziehungen zueinander untersucht.

Zufallsvariablen können dabei beobachtet oder unbeobachtet sein. Dabei sind nur beobachtete Zufallsvariablen in Experimenten bekannt. Unbeobachtete Zufallsvariablen hingegen sind während der Experimente nicht bekannt. Es gibt keine Möglichkeit, die

unbeobachteten Zufallsvariablen direkt abzulesen. Um dennoch Aussagen über sie zu treffen, kann oft der Satz von Bayes ausgenutzt werden. Von sicherem oder unsicherem Wissen auf Aussagen über unbekannte Sachverhalte zu schließen, wird als Inferenz bezeichnet. Sie unterteilt sich in diagnostische und kausale Unterstützung. [35]

1. **Kausale Unterstützung** beschreibt die Beobachtung einer Ursache der Unbekannten. Ist die Ursache bekannt, kann auf deren Wirkung geschlossen werden.
2. **Diagnostische Unterstützung** funktioniert genau umgekehrt. Um Aussagen über eine unbekannte Ursache zu treffen, ist deren beobachtete Wirkung hilfreich. Da die Wirkung aus bestimmten Ursachen folgt, gibt sie Aufschluss über die Unbekannte.

3.2 Bayes'sche Netze

Bayes'sche Netze sind graphische Modelle für probabilistische Beziehungen zwischen einer Menge von Zufallsvariablen. Sie bieten verschiedene Vorteile. So sind sie in der Lage, unvollständige Datensätze zu verarbeiten, liefern Informationen über kausale Beziehungen und vereinfachen die Kombination von Fachwissen und Daten. [20] Dieser Abschnitt ist eine kurze Einführung in die Theorie Bayes'scher Netze und folgt dabei den Bezeichnungen und Definitionen in [11], [20] und [35].

Zunächst werden einige graphentheoretische Begriffe aus [11] und [35] eingeführt:

Definition 5 (Graph).

Ein **Graph** $G(V, E)$ besteht aus einer Menge von **Knoten** $V = \{V_1, V_2, \dots\}$ und einer Menge von **Kanten** $E = \{E_1, E_2, \dots\}$. Dabei hat jede Kante zwei Endpunkte, die Knoten und nicht notwendigerweise verschieden sind.

Die Kanten, die zwei Knoten V_1 und V_2 verbinden, bilden einen **Pfad** von V_1 nach V_2 . Sind Anfangs- und Endknoten eines Pfades identisch, so ist dieser ein **Zyklus**.

Ein Graph heißt **gerichtet**, wenn das Paar der Endpunkte geordnet ist. Sind (V_1, V_2) Endpunkte der Kante E_1 , lässt sich schreiben:

$$V_1 \xrightarrow{E_1} V_2. \quad (3.3)$$

Ein gerichteter **azyklischer** Graph enthält keinen gerichteten Zyklus.

Seien V_1 und V_2 Knoten in einem gerichteten azyklischen Graphen $\mathbb{G} = (V, E)$. V_1 ist ein **Elternelement** von V_2 , wenn es eine Kante von V_1 nach V_2 gibt. Analog ist V_2 ein **Kinderelement** von V_1 . Existiert ein Pfad von V_1 nach V_2 , so ist V_1 Vorfahre von V_2 und V_2 dessen Nachfahre.

Um Bayes'sche Netze zu definieren, wird die Markov-Bedingung benötigt. Sie identifiziert die Knoten eines gerichteten Graphen mit Zufallsvariablen

Nach [35] gilt:

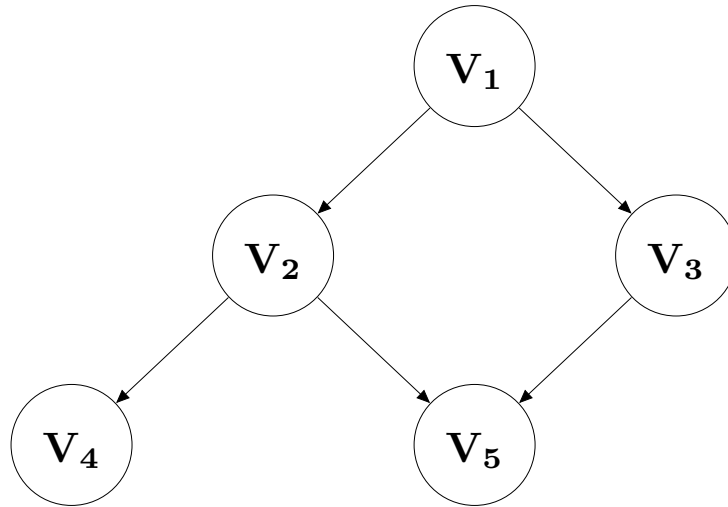


Abbildung 3.1: Beispiel für einen gerichteten Graphen mit fünf Knoten. Die Kanten stellen Ursache-Wirkungs-Beziehungen dar. Der Graph ist azyklisch.

Definition 6 (Markov-Bedingung).

Sei V eine Menge von Zufallsvariablen und sei P deren gemeinsame Wahrscheinlichkeitsverteilung. Weiter sei E eine Menge von Kanten und sei $\mathbb{G} = (V, E)$ ein gerichteter azyklischer Graph.

(\mathbb{G}, P) erfüllt die Markov-Bedingung, wenn für jede Zufallsvariable $X \in V$ gilt: $\{X\}$ ist bedingt unabhängig von der Menge seiner Nichtnachfahren unter der Menge seiner Elternelemente.

Ist X ein Wurzelknoten, so hat er keine Elternelemente. Gilt die Markov-Bedingung, ist er unabhängig von allen Nichtnachfahren.

Im Graphen in Abbildung 3.1 existiert nur ein Wurzelknoten V_1 . Alle anderen Knoten im Graphen sind Nachfahren von V_1 . Für ihn ist die Markov-Bedingung trivialerweise erfüllt. Für die anderen Knoten muss gelten:

- $\{V_2\}$ ist unabhängig von $\{V_3\}$ unter der Bedingung $\{V_1\}$,
- $\{V_3\}$ ist unabhängig von $\{V_2, V_4\}$ unter der Bedingung $\{V_1\}$,
- $\{V_4\}$ ist unabhängig von $\{V_1, V_3, V_5\}$ unter der Bedingung $\{V_2\}$,
- $\{V_5\}$ ist unabhängig von $\{V_1, V_4\}$ unter der Bedingung $\{V_2, V_3\}$.

Mit der Definition gerichteter Graphen und der Markov-Bedingung lassen sich Bayes'sche Netze wie folgt definieren:

Definition 7 (Bayes'sches Netz).

Sei V eine Menge von Zufallsvariablen und sei P deren gemeinsame Wahrscheinlich-

keitsverteilung. Weiter sei E eine Menge von Kanten und sei $\mathbb{G} = (V, E)$ ein gerichteter azyklischer Graph. Dann ist (\mathbb{G}, P) ein Bayes'sches Netz, wenn es die Markov-Bedingung erfüllt.

Ist die Markov-Bedingung für Abbildung 3.1 erfüllt, so repräsentiert der Graph ein Bayes'sches Netz.

Mit Hilfe Bayes'scher Netze kann unsicheres Wissen repräsentiert und mit Expertenwissen verknüpft werden [25]. Um ein Bayes'sches Netz zu modellieren, sind die folgenden Schritte auszuführen:

Eine konkrete Fragestellung muss benannt werden, anhand welcher Zielgrößen festgelegt werden. Dann müssen Zufallsvariablen identifiziert werden, die relevant für die zu bestimmende Zielgröße sind. Diese Zufallsvariablen werden als Knoten dargestellt. Weiter werden gerichtete Kanten hinzugefügt, die kausale Zusammenhänge und Abhängigkeiten zwischen diesen Hypothesen darstellen. Eine gerichtete Kante führt dabei vom ursächlichen Knoten zum Wirkungsknoten. Ist weiter eine Wahrscheinlichkeitsverteilung für die Zufallsvariablen des Graphen gegeben und die Markov-Bedingung erfüllt, so ist ein Bayes'sches Netz konstruiert. Es gelten die im folgenden Abschnitt beschriebenen Beobachtungen.

3.3 Wahrscheinlichkeiten in Bayes'schen Netzen

Ist die Struktur des Bayes'schen Netzes festgelegt, so kann es mit Wahrscheinlichkeiten quantifiziert werden. In diesem Abschnitt soll zunächst erklärt werden, wie *Lernen* in Bayes'schen Netzen funktioniert und das Netz mit Wahrscheinlichkeiten befüllt wird. Dann wird gezeigt, wie ein mit Daten versehenes Netz für Inferenzen über unbekannte Zufallsvariablen genutzt werden kann.

Ein Bayes'sches Netz besteht aus Knoten, die die Wahrscheinlichkeit des Eintretens gewisser Hypothesen repräsentieren. Damit es für Inferenzen genutzt werden kann, müssen die Hypothesen mit zugehörigen Wahrscheinlichkeiten versehen werden. Hierfür kann auf Expertenwissen oder auf Evidenzen aus bereits durchgeführten Experimenten zurückgegriffen werden.

Dabei werden A-priori- und A-posteriori-Wahrscheinlichkeiten unterschieden. Als A-priori-Wahrscheinlichkeit werden Wahrscheinlichkeiten bezeichnet, die die Grundwahrscheinlichkeit eines Ereignisses angeben. Sie sind Wahrscheinlichkeiten, die vor Inferenzbetrachtungen bekannt sind. Analog bezeichnen A-posteriori-Wahrscheinlichkeiten diejenigen Wahrscheinlichkeiten, die nach Inferenzschlüssen aus Vorwissen und Evidenzen entstehen. [35]

Sollen A-priori-Wahrscheinlichkeiten experimentell bestimmt werden, die Histogramm-Methode verwendet werden. Hierfür wird der Wertebereich einer Zufallsvariable in mehrere Intervalle aufgeteilt, die jeweils eine Klasse bilden. In Experimenten wird die Zufallsvariable

ble beobachtet und ihre diskreten Werte gespeichert. Durch Zählen relativer Häufigkeiten kann so die diskrete A-priori-Wahrscheinlichkeit bestimmt werden. [9]

Für Knoten ohne Vorfahren können oft direkt Wahrscheinlichkeiten angegeben werden. Hat ein Knoten Vorfahren, so werden bedingte Wahrscheinlichkeiten angegeben, die beziffern, wie wahrscheinlich die Hypothese in Abhängigkeit der Hypothesen der Vorgängerknoten ist. Der Prozess des Bestimmens der Wahrscheinlichkeiten wird als *Lernen* im Bayes'schen Netz bezeichnet. [9], [25]

Auch zur Bestimmung bedingter Wahrscheinlichkeiten kann die Histogramm-Methode angewandt werden. Soll eine Näherung für $P(V|W_1, \dots, W_n)$ bestimmt werden, müssen dafür alle Zufallsvariablen V, W_1, \dots, W_n beobachtet sein. Weiter müssen Diskretisierungen für die Zufallsvariablen festgelegt werden. Dann kann jede Beobachtung den entsprechenden diskreten Klassen zugeordnet werden und die Häufigkeiten für (v, w_1, \dots, w_n) für alle $v \in V, w_1 \in W_1, \dots, w_n \in W_n$ festgestellt werden. Anschließend können die relativen Häufigkeiten bestimmt werden. Sei mit N_v die Anzahl des Auftretens von (v, w_1, \dots, w_n) bezeichnet. Dafür muss die Bedingung

$$\sum_{v \in V} P(v|w_1, \dots, w_n) = 1 \text{ für alle } w_1 \in W_1, \dots, w_n \in W_n$$

erfüllt sein. Insgesamt gilt

$$P(v|w_1, \dots, w_n) \approx \frac{N_v}{\sum_{u \in V} N_u}.$$

Um solche bedingten Wahrscheinlichkeiten anzugeben, werden **Verbindungsmatrizen** aufgestellt, die die Wahrscheinlichkeitswerte für jede Kombination aus möglichen Ereignissen der Vorgängerknoten und des Knotens selbst darstellen. Sie erleichtern das Rechnen in Bayes'schen Netzen.

Definition 8 (Verbindungsmatrix).

Der Knoten K habe n Vorgängerknoten V^1, \dots, V^n . Dabei stelle $V_{j_1}^1, \dots, V_{j_n}^n$ beliebige, aber feste Hypothesen dar, die in V^1, \dots, V^n repräsentiert werden. Mit K_i für $i = 1, \dots, h, h \in \mathbb{N}$, seien die in K repräsentierten Hypothesen bezeichnet. Dann ist die Verbindungsmatrix $M(K|V^1, \dots, V^n)$ des Knotens K gegeben durch

$$(M(K|V^1, \dots, V^n))_{i, j_1, \dots, j_n} = P(K_i|V_{j_1}^1, \dots, V_{j_n}^n).$$

Die Verbindungsmatrizen erfüllen stets die Bedingung für $h \in \mathbb{N}$ in K repräsentierte Hypothesen

$$\sum_{i=1}^h (M(K|V^1, \dots, V^n))_{i, j_1, \dots, j_n} = 1.$$

Die Dimension der Verbindungsmatrix von K ist dabei immer um 1 größer als die Anzahl der Vorgängerknoten von K . Stochastische Matrizen sind dadurch definiert, dass ihre Spalten- oder Zeilensumme 1 ergibt. Folglich sind Verbindungsmatrizen stochastisch.

Die folgende aus [26] übernommene Kettenregel ist eine wichtige Eigenschaft Bayes'scher Netze. Sie vereinfacht die Berechnung der Wahrscheinlichkeiten in einem Bayes'schen Netz maßgeblich.

Satz 2 (Kettenregel für Bayes'sche Netze).

(\mathbb{G}, P) erfülle die Markov-Bedingung. Sei $P(V)$ die gemeinsame Wahrscheinlichkeitsverteilung der Zufallsvariablen $V = \{V_1, V_2, \dots, V_n\}$ und bezeichne $X_{pa(V_i)}$ die Menge aller Elternelemente von V_i für $1 \leq i \leq n$. Dann gilt

$$P(V) = \prod_{1 \leq i \leq n} P(V_i | X_{pa(V_i)}). \quad (3.4)$$

Die Kettenregel ergibt sich direkt aus der Markov-Bedingung.

Bayes'sche Netze sind für Inferenzen nutzbar. Dafür können diagnostische und kausale Unterstützung wie in Abschnitt 3.2 beschrieben ausgenutzt werden. Experimente können oft durchgeführt werden, um Evidenzen zu bestimmen. Dank des Satz von Bayes lassen sich aus den Beobachtungen und dem quantifizierten Bayes'schen Netz A-posteriori-Verteilungen bestimmen. Hierfür kann der Maximum-A-posteriori-Schätzer verwendet werden.

Sei mit Kettenregel und Satz von Bayes die Wahrscheinlichkeit $P(V|W)$ für unbekanntes V und observierbare Zufallsvariablen W bestimmt. Um auf Grundlage dieser Evidenzen auf die wahrscheinlichste Hypothese zu schließen, wird ein Maximum-A-posteriori-Schätzer (MAP) verwendet. Dabei wird die Realisierung der Zufallsvariable V geschätzt. Vorausgesetzt wird, dass die A-priori-Wahrscheinlichkeiten $P(\theta|W)$ für alle Realisierungen $X = \theta$ bekannt sind. Der MAP liefert dann für die beobachteten Evidenzen diejenige Realisierung von X , die am wahrscheinlichsten ist. [20], [35]

Definition 9 (Maximum-A-posteriori-Schätzer).

Der Maximum-A-posteriori-Schätzer - kurz MAP - bezeichnet denjenigen Wert der Parameter θ , der die beobachteten Daten W am wahrscheinlichsten macht. Er ist definiert als

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} P(\theta|W). \quad (3.5)$$

Mit dem MAP können im Bayes'schen Netz unbekannte Wahrscheinlichkeiten geschätzt werden. Dabei wird die A-priori-Wahrscheinlichkeit lediglich als Normierungsfaktor genutzt, die den Schätzwert nicht verändert. [35]

3.4 Datenfusion

Datenfusion beschreibt die Integration von Informationen aus mehreren Datenquellen. Ziel dabei ist, spezifische zusammengeführte Daten über eine Einheit zu erhalten. Was Gegenstand der Datenfusion ist und als Einheit verstanden wird, ist kontext- und anwendungsabhängig. [19]

Die Multisensor-Datenfusion bezeichnet die Fusion von Daten, die von unterschiedlichen Sensoren stammen. Sie hat in den letzten Jahren mehr und mehr Aufmerksamkeit erhalten und umfasst eine Vielzahl unterschiedlicher Techniken. Dabei ist die Kombination von Daten keine neue Idee. Menschen und Tiere kombinieren seit jeher unterschiedliche Sinneseindrücke zu gemeinsamen Ergebnissen. Oft werden beispielsweise akustische und visuelle Signale verknüpft. Ist eine Stelle nicht gut einsehbar, so nutzen Mensch und Tier den Hörsinn, um etwaige Gefahren zu erkennen. [19]

Mit dem technologischen Fortschritt steigt die Anzahl verfügbarer Sensoren, die Daten erfassbar machen. Dies erfordert neue Techniken, Daten zu verarbeiten und zu fusionieren.

Multisensor-Datenfusion wird beispielsweise im Bereich des Automatisierten Fahrens genutzt. Voraussetzung für sicheres Fahren bei hoher Automatisierungsstufe ist, dass die Verkehrssituation zuverlässig erkannt wird. Dazu wird ein Umgebungsmodell erstellt, das Informationen zur Situationsanalyse enthält. Es ist eine Abbildung der tatsächlichen Umgebung, die Verkehrsteilnehmer, Infrastruktur und weitere relevante Informationen enthält. Auf Grundlage dieses Modells können dann automatisiert Entscheidungen getroffen werden. Das Modell muss die Umwelt dafür genau abbilden. So müssen Fußgänger nicht nur erkannt werden, sondern es muss auch interpretiert werden, ob diese die Straße überqueren und Abbremsen oder Ausweichmanöver nötig sind. Um ein Umgebungsmodell zu erstellen, wird Datenfusion eingesetzt. Abbildung 3.2 zeigt, wie verschiedene Sensoren wie Kameras, Radar, LIDAR und GPS zur Erstellung eines Umgebungsmodells fusioniert werden. [17]

Dieses Modell besteht aus der Erkennung einzelner Objekte, die zu einer Darstellung der gesamten Umgebung zusammengefügt werden. So kann beispielsweise Fußgängererkennung durchgeführt werden. Die einzelnen Sensoren zeichnen sich dabei durch unterschiedliche Stärken und Schwächen aus. Erst ihre Kombination mittels Datenfusion führt zu zuverlässigen Ergebnissen. Redundanz kann darüber hinaus genutzt werden, um die erfassten Informationen zu verifizieren. [17]

Datenfusion wird auch eingesetzt, um Probleme in der Robotik, der automatisierten Kontrolle industrieller Fertigungssysteme oder der medizinischen Anwendung zu lösen [19]. Weiter kann Datenfusion genutzt werden, um Verkehr zu erfassen. In [23] wird ein solcher fusionsbasierter Ansatz verfolgt, um Verkehrszustände durch Audio- und Bilddaten zu erfassen.

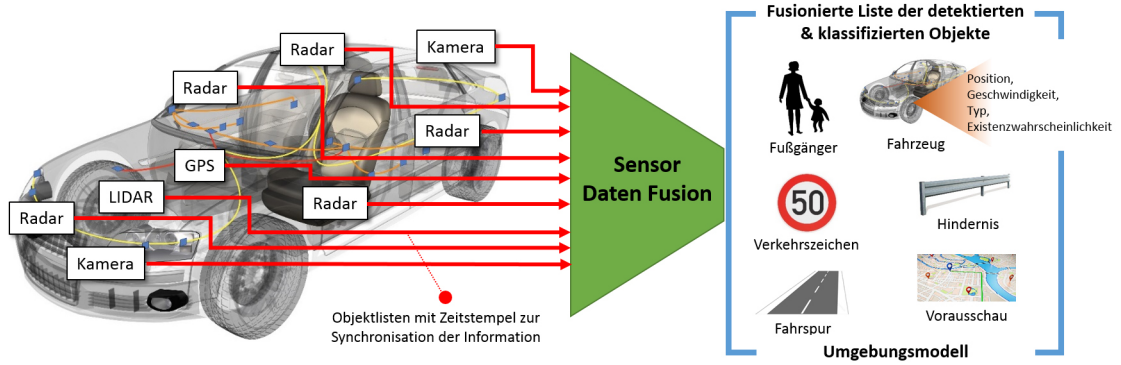


Abbildung 3.2: Darstellung verschiedener Sensoren, die im Automatisierten Fahren genutzt werden, und durch Datenfusion zugänglich gemachte Informationen über verschiedene Objekte, aus [17].

Mathematisch kann Datenfusion definiert werden mit dem Funktional ζ , das die Verfahren, die zur Datenfusion genutzt werden, repräsentiert. Dann ist die Fusion beschrieben durch

$$x \approx \hat{x} = \zeta(z_{(1)}, \dots, z_{(n)}). \quad (3.6)$$

Dabei ist \hat{x} das Fusionsergebnis, das der optimalen Schätzung von x entspricht. Wie gut diese Schätzung ist, hängt davon ab, wie zuverlässig die einzelnen Daten $z_{(i)}$ für $i = 1, \dots, n$ sind und wie gut die Verfahren zur Datenfusion ζ sind. [27]

Auch Bayes'sche Netze können zur Datenfusion verwendet werden. Hierfür werden die Eingangsdaten als beobachtete Zufallsvariablen definiert. In einem Bayes'schen Netz werden sie modelliert. Die zu schätzende Zielgröße ist dabei das Ergebnis der Datenfusion und muss ebenfalls als Zufallsvariable ins Modell integriert werden. Sie ist dabei eine Ursache der Detektion durch alle Sensoren.

Es kann etwa ein Bayes'sches Netz aufgestellt werden, um aus den Bluetooth- und C2X-Geschwindigkeitsdetektionen wie in den Abschnitten 2.2.1 und 2.2.2 die tatsächliche Geschwindigkeit zu schätzen. In Abbildung 3.3 wurde ein Bayes'sches Netz modelliert, das die unbekannte tatsächliche Geschwindigkeit und die Detektionsgeschwindigkeiten V_{BT} und V_{C2X} enthält.

V ist dabei ursächlich für die Messgeschwindigkeiten V_{BT} und V_{C2X} . Das Netz kann genutzt werden, um von den Sensordaten aus C2X- und Bluetooth-Detektion mittels Datenfusion auf die tatsächliche Gegebenheit zu schließen. In diesem Fall beschreibt in Gleichung 3.6 ζ den Vorgang, die Wahrscheinlichkeitsverteilungen zu quantifizieren und

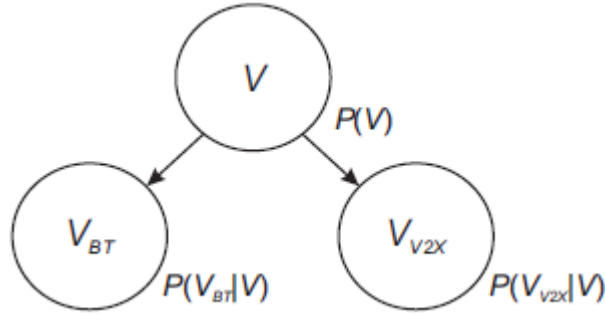


Abbildung 3.3: Darstellung von Ziel- und Hilfsgrößen im Bayes'schen Netz. Die unbekannte mittlere Momentangeschwindigkeit V ist die Ursache von Bluetooth- und C2X-Detektion V_{BT} bzw. V_{V2X} , aus [24].

mittels eines Schätzers die Schätzung der Geschwindigkeit zu bestimmen. Es gilt:

$$V \approx \hat{V} = \zeta(V_{BT}, V_{C2X}).$$

3.5 Zugrunde liegendes Modell

In [24] wurde ein Modell zur Verkehrszustandsschätzung durch ein Bayes'sches Netz erstellt. Ausgehend von den dort beschriebenen Ergebnissen soll das Modell überarbeitet werden. Hier wird ein Überblick über das Modell gegeben, das im folgenden Kapitel weiterentwickelt wird.

Die Autoren befassen sich mit der Frage, wie sich C2X-Kommunikation für eine Verkehrszustandsschätzung nutzen lässt, selbst wenn sie nur in einem Teil der Fahrzeuge verfügbar ist. Dabei verfolgen sie den Ansatz, ein Bayes'sches Netz zur Erfassung des Verkehrszustands zu modellieren.

Zur Unterstützung soll ein weiterer Sensor hinzugezogen werden, der Ergebnisse für eine größere Zahl von Fahrzeugen liefert. Die Wahl fällt auf ein Bluetooth-basiertes Verfahren, wie in 2.2.2 beschrieben.

Zunächst ist dabei zu klären, anhand welcher Größe Aussagen über den Verkehrszustand gemacht werden sollen. Wichtig dabei ist, dass sie gleichermaßen aussagekräftig und gut erfassbar ist. Während mittels C2X-Kommunikation eine Vielzahl von Daten übertragbar und für die Modellierung verwendbar ist, kommt der in 2.2.2 beschriebene Ansatz zur Geschwindigkeitsdetektion mittels Bluetooth nur mit den eindeutigen IDs der Bluetooth-Sender und dem zeitlichen Abstand der Detektionen aus. Daraus lässt sich eine Schätzgeschwindigkeit berechnen. Da die Bluetooth-Detektion die mittlere Momentangeschwindigkeit als Zielgröße hat, ist es naheliegend, auch Geschwindigkeiten

als Zielgröße für C2X und Datenfusion zu verwenden. Daher wird auch die mit C2X detektierte Geschwindigkeit als Hauptbezugsdatum der C2X-Detektion gewählt.

Die Zielgröße des Bayes'schen Netzes ist daher die mittlere Momentangeschwindigkeit v über dem gemeinsamen Bluetooth- und C2X-Detektionsbereich. Hierfür werden die Detektoren so angeordnet, dass der C2X-Bereich den gesamten Bluetooth-Detektionsbereich einschließt. So können Untersuchungen auf dem gemeinsamen Bereich durchgeführt werden. Dies ist in Abbildung 3.4 dargestellt.

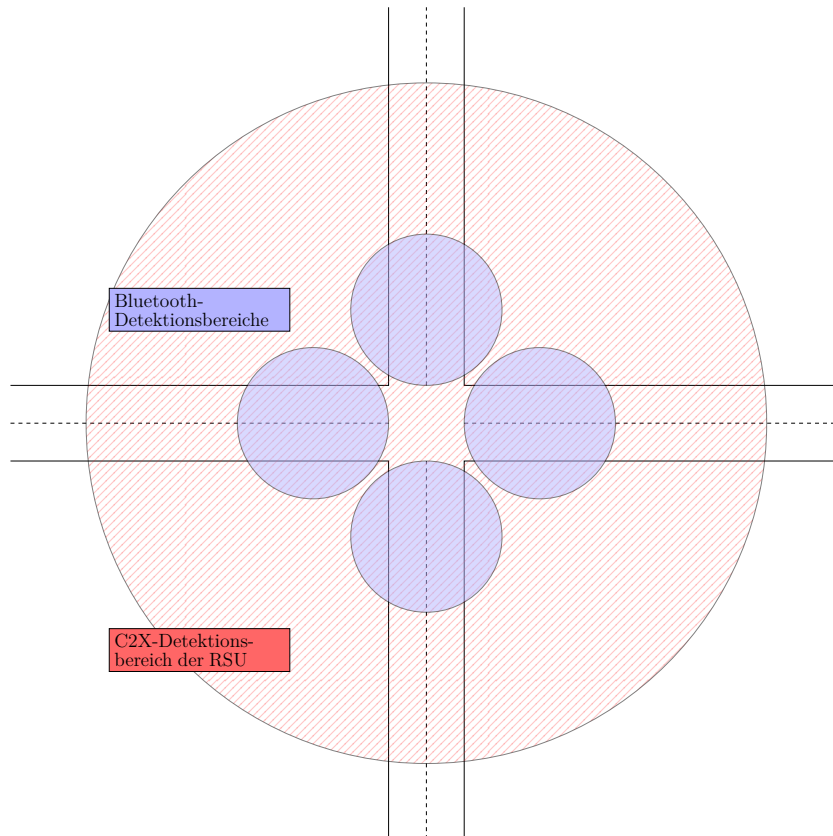


Abbildung 3.4: Anordnung der Bluetooth- und C2X-Detektionsbereiche am Knotenpunkt.

Zunächst wurde für die Fusion ein naives Modell aufgestellt, dass lediglich die tatsächliche Geschwindigkeit, die Bluetooth-Detektion und die C2X-Detektion umfasst. Die Größen wurden als Zufallsvariablen definiert und ihre Kausalitäten dargestellt. Die tatsächliche mittlere Momentangeschwindigkeit v ist dabei ursächlich für die beiden Messgeschwindigkeiten für Bluetooth v_{BT} und C2X v_{C2X} . Letztere sind untereinander bedingt unabhängig.

Repräsentiert durch ein Bayes'sches Netz stellt sich dies wie in Abbildung 3.3 dar.

Im naiven Modell werden die Detektionsgeschwindigkeiten mit der zu schätzenden tatsächlichen Geschwindigkeit verknüpft. Außer Acht bleiben jedoch andere Faktoren, die den Verkehrsfluss und die Geschwindigkeit beeinflussen. Um die Ergebnisse des Modells zu verbessern, bietet es sich an, zusätzliche Verkehrskenngrößen und beeinflussende Faktoren in die Schätzung zu integrieren. Als relevant identifiziert wurden im nächsten Schritt die folgenden Zusammenhänge [24]:

- Direkten Einfluss auf die Geschwindigkeit V nehmen die Lichtsignalanlage TLC (traffic light control) und die Verkehrsdichte D . Dabei ist zu beachten, dass auch untereinander Einflüsse bestehen. So besteht eine Abhängigkeit der Lichtsignalanlage (LSA) von der Verkehrsdichte, wenn die LSA verkehrsabhängig gesteuert ist. Bei einer Festzeitsteuerung kann sogar von einer Kausalität in die andere Richtung gesprochen werden. Im finalen Netz müssen solche Abhängigkeiten genau geprüft und an den Anwendungsfall angepasst werden.
- Die Bluetooth-Detektionen sind keine Geschwindigkeitsdetektionen im eigentlichen Sinne. Vielmehr ist der Bluetooth-Sender in der Lage, Aussagen über die Belegung des Erfassungsbereichs zu machen. Daher wird die Belegung O als ursächlich für die Bluetooth-Detektion angenommen. Die Belegung selbst ist wiederum von Verkehrsdichte D und tatsächlicher Geschwindigkeit V abhängig.
- Um die naturgemäß fehlerbehaftete Bluetooth-Detektion V_{BT} zu verbessern, sollen zusätzliche Einflussfaktoren abgebildet werden. So wird die letzte mit Bluetooth detektierte Geschwindigkeit V'_{BT} als relevant für die aktuelle Bluetooth-Detektion angenommen. Auch die Geschwindigkeitsdifferenz ΔV zwischen diesen Detektionen wird als Einfluss angenommen. ΔV wird seinerseits von der aus der Detektion resultierende geschätzte Zeitlücke ΔT beeinflusst.

Der Graph, der die beschriebenen Kausalitäten darstellt, ist in Abbildung 3.5 dargestellt. Ein Bayes'sches Netz auf Basis dieses Graphens aufzustellen ist prinzipiell möglich, wirft aber verschiedene Probleme auf:

- Zunächst ist das Netz mehrfachverbunden. Dies bedeutet, dass es von einem Knoten zu einem anderen mehrere Pfade gibt [35]. Das ist zum Beispiel für die Knoten D und V_{BT} der Fall. Es existieren die Pfade $D-O-V_{BT}$, $D-V-V_{BT}$, $D-TLC-V-V_{BT}$, $D-V-O-V_{BT}$ und $D-TLC-V-O-V_{BT}$. Mehrfachverbundenheit in Bayes'schen Netzen sorgt für eine komplexere Inferenz. Daher ist ein einfach verbundenes Bayes'sches Netz vorzuziehen, sofern dies die dargestellten Zusammenhänge weiterhin abbildet.
- Darüber hinaus müssen die verwendeten Größen auch tatsächlich messbar sein. Das ist etwa im Falle der Belegung problematisch. Der Bluetooth-Sender ist in der Lage, die mit Bluetooth ausgestatteten Fahrzeuge zu erkennen. Alle Fahrzeuge ohne Bluetooth-Ausstattung fehlen jedoch. Ist der Prozentsatz der Bluetooth-ausgestatteten Fahrzeuge p_{BT} bekannt, kann mit der Anzahl der detektierten

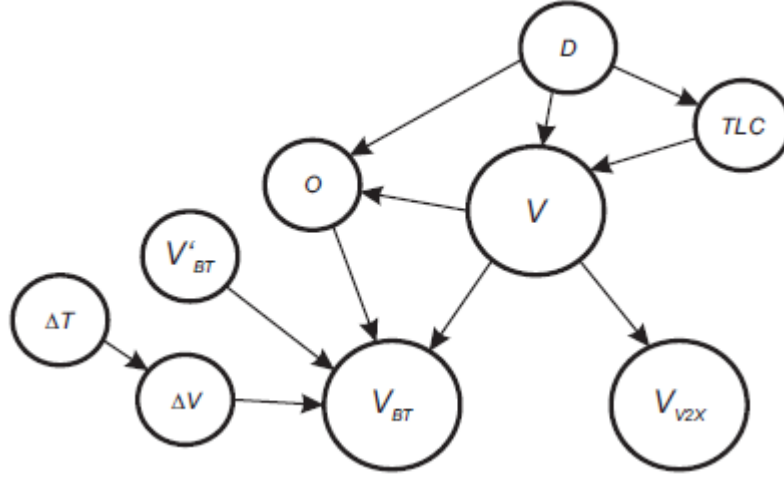


Abbildung 3.5: Identifizierte Abhängigkeiten verschiedener Verkehrskenngrößen und Detektordaten, aus [24].

Fahrzeuge $n_{BT}(t)$ und einem Korrekturfaktor $c(t)$ die Gesamtzahl der Fahrzeuge $n(t)$ zum Zeitpunkt t geschätzt werden wie in [24]:

$$n(t) \approx \frac{n_{BT}(t)}{p_{BT}} c(t).$$

Diese unterliegt jedoch großen Schwankungen. Von der Integration nur näherungsweise messbarer Einflussparameter soll daher abgesehen werden.

Daher werden die Knoten wieder reduziert. In Abbildung 3.6 ist dies dargestellt. Dabei wurde verzichtet auf:

- O , D : Die Knoten sind in realistischen Anwendungen schwer exakt erfassbar und sollen zur realitätsnahen Modellierung nicht verwendet werden.
- TLC : Der Knoten hat direkte Auswirkungen auf die zu schätzende Geschwindigkeit V und soll zur Vereinfachung weggelassen werden.
- ΔT : Die Informationen aus ΔT und ΔV sind redundant. Daher kann auf einen der Knoten verzichtet werden.

Der gerichtete Graph kann nun experimentell mit Wahrscheinlichkeitsverteilungen befüllt werden. Ein Bayes'sches Netz entsteht.

Nachdem ein Lernalgorithmus zur Quantifizierung des Bayes'schen Netzes verwendet wurde, konnte die Datenfusion simulativ im Verkehrssimulator SUMO durchgeführt werden. Für die simulativen Untersuchungen wurde das in Abschnitt 2.3.3 beschriebene

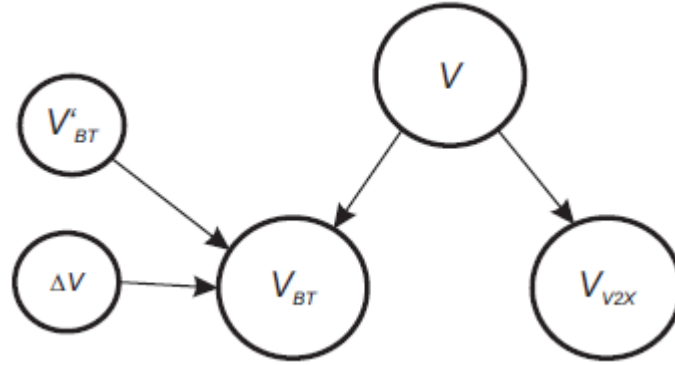


Abbildung 3.6: Gerichteter Graph für das reduzierte Bayes'sche Netz. Neben der Zielgeschwindigkeit V und den Detektionen V_{BT} und V_{V2X} sollen die letzte Bluetooth-Detektion V'_{BT} sowie die detektierte Geschwindigkeitsdifferenz ΔV zur Verbesserung des Fusionsergebnisses verwendet werden, aus [24].

Szenario RiLSA verwendet. Dabei unterscheidet sich lediglich die Verkehrsnachfrage. In Abhängigkeit der C2X-Ausstattungsrate wurden die Ergebnisse in 3.7 erzielt.

V2X p. rate	v_{RMS} [m/s] and q_c [%]			
	north	east	south	west
1%	2.3 (34.6)	4.7 (33.6)	1.7 (37.3)	5.3 (38.6)
2%	2.3 (37.2)	4.8 (36.0)	1.8 (43.3)	4.9 (41.9)
5%	2.3 (42.0)	4.2 (40.5)	2.0 (49.8)	4.5 (47.3)
10%	2.4 (51.9)	3.9 (47.4)	2.1 (52.6)	3.9 (56.6)
20%	2.3 (62.3)	3.4 (61.9)	2.2 (67.6)	3.2 (69.5)
50%	2.2 (84.8)	2.6 (85.7)	2.1 (88.7)	2.6 (90.9)
100%	1.8 (99.9)	2.1 (100.0)	1.6 (99.9)	2.1 (100.0)

Abbildung 3.7: Ergebnisse der Schätzung aus [24]. Genauigkeit als Root Mean Square Error v_{RMS} und Vollständigkeit q_c der Schätzung in Abhängigkeit von C2X-Penetrationsrate und Kreuzungsarm.

Es ist zu sehen, dass sich der Root-Mean-Square-Error v_{RMS} bei steigender C2X-Ausstattungsrate verringert. Für die Ost-West-Achse, auf der eine höhere Verkehrsnachfrage als in Nord-Süd-Richtung besteht, fällt der Fehler deutlich von 4,7 bzw. 5,3 m/s auf je 2,1 m/s, wenn sich die Ausstattungsrate schrittweise von 1% auf 100% erhöht. In Nord-Süd-Richtung ist v_{RMS} schon für niedrige Ausstattungsgrade klein.

Die Vollständigkeit q_c ist ein Maß dafür, ob Daten für alle Zeitpunkte vorliegen, an denen Daten vorliegen können. Sie gibt prozentual an, zu wie vielen Zeitpunkten eine Schätzung der Geschwindigkeit vorliegt. Sie erhöht sich für alle Kreuzungsarme mit der

Erhöhung des C2X-Ausstattungsgrades. Sie liegt für 1% Ausstattung noch bei 30 - 40%, erreicht aber schon bei 50% Ausstattung in allen Richtungen über 80%. Gerade bei niedrigen Ausstattungen mit C2X fällt hier die Bluetooth-Ausstattung von 30% stark ins Gewicht.

Dieses Modell soll im Folgenden als Grundlage genommen werden, um eine Verkehrszustandsschätzung unter Ausnutzung von C2X- und Bluetooth-Detektionen zu ermöglichen. Dazu wird das bestehende Modell evaluiert und dessen Einflussparameter kritisch betrachtet. Auf der Basis der Untersuchungen soll ein neues Bayes'sches Netz modelliert werden. Ziel dabei ist, die Ergebnisse aus [24] zu verbessern. Die Ergebnisse in Abbildung 3.7 sollen dabei als Referenzwerte dienen.

4 Modell und Implementierung

In diesem Kapitel wird die mathematische Modellierung der Verkehrszustandsschätzung beschrieben. Basierend auf der im vorigen Kapitel vorgestellten Methodik und den dort erklärten Kenngrößen wird ein Bayes'sches Netz entwickelt, das eine Schätzung der mittleren Momentangeschwindigkeit ermöglicht. Dazu wird das in Abschnitt 3.5 beschriebene Modell analysiert und erweitert. Im ersten Abschnitt 4.1 wird diese Modellierung vollzogen. Benötigte Zufallsvariablen werden identifiziert und kausale Zusammenhänge abgebildet. Anschließend sind die Variablen geeignet zu diskretisieren, sodass das Bayes'sche Netz quantifiziert werden kann. Es folgt Abschnitt 4.2, der die Implementierung des Modells inklusive der Quantifizierung und des Einsatzes zur Datenfusion beschreibt. Damit ist der Modellaufbau gegeben, der im folgenden Kapitel für experimentelle Untersuchungen genutzt wird.

4.1 Modellierung des Bayes'schen Netzes

In diesem Abschnitt wird das verwendete Modell zur Verkehrszustandsschätzung beschrieben. Dazu wird ein Bayes'sches Netz basierend auf dem in 3.5 vorgestellten Modell aufgestellt. Es wird erklärt, wie das Netz modelliert wird und welche Verkehrskenngrößen ausgenutzt werden. Dabei werden Knoten und deren kausale Zusammenhänge für das Netz identifiziert. Anschließend wird eine geeignete Diskretisierung der Größen im Netz vorgenommen.

Bei einer Verkehrszustandsschätzung können Kenngrößen - wie etwa Verkehrsstärke und Geschwindigkeit - identifiziert werden, zwischen denen ein kausaler Zusammenhang besteht, dessen explizite Quantifizierung sich aber schwierig gestaltet. Am Beispiel von Verkehrsstärke, -dichte und Geschwindigkeit lässt sich dies gut veranschaulichen.

Für das räumlich-zeitliche Mittel von Geschwindigkeit v_{RZ} , Verkehrsdichte D_{RZ} und -stärke q_{RZ} lässt sich die Kontinuitätsgleichung des Straßenverkehrs wie in [41] aufstellen:

$$q_{RZ} = v_{RZ} D_{RZ}. \quad (4.1)$$

Werden jedoch lokale oder momentan ermittelte Messgrößen benötigt, so ähneln deren Zusammenhänge den räumlich-zeitlichen Mitteln nur, wenn die baulichen Verhältnisse

einen ähnlichen Verkehrsablauf für eine längere Strecke ermöglichen. Dies ist gerade an Knotenpunkten eine Voraussetzung, die in der Regel nicht erfüllt werden kann. Kommen weitere Veränderungen wie Geschwindigkeitsbeschränkungen, Fußgängerquerverkehr oder Hindernisse hinzu, verändert sich diese Beziehung noch weiter und die analytische Formulierung des Zusammenhangs wird komplizierter. [41]

Im untersuchten Szenario kommt zudem noch die Phase der Lichtsignalanlage als regulierender Faktor hinzu, der zu noch komplexeren Zusammenhängen führt. Die Modellierung mit einem Bayes'schen Netz bietet den Vorteil, dass auf eine explizite analytische Darstellung der Zusammenhänge verzichtet werden kann.

Es genügt, die Einflussfaktoren als Zufallsvariablen zu definieren und deren kausale Zusammenhänge zu untersuchen. Dabei muss jedoch mit besonderer Vorsicht vorgegangen werden. Schließlich muss nach [35] in einem Bayes'schen Netz sichergestellt sein, dass *alle* kausalen Einflussfaktoren in der richtigen Granularität dargestellt werden. Dies ist faktisch unmöglich, da eine Vielzahl von Faktoren wie Witterungsbedingungen, Ablenkung der Verkehrsteilnehmer oder der Straßenzustand einen Einfluss auf den Verkehr haben können. Das hier entstehende Bayes'sche Netz kann also nur als Annäherung an den Verkehrszustand verstanden werden und muss seine Anwendbarkeit noch quantitativ legitimieren.

Soll ein Bayes'sches Netz aufgestellt werden, muss zunächst eine Festlegung auf alle Zufallsvariablen erfolgen, die ins Modell aufgenommen werden sollen. Dann müssen die Ursache-Wirkungs-Beziehungen zwischen allen Zufallsvariablen identifiziert werden. Ist die strukturelle Modellierung des Netzes abgeschlossen, kann die probabilistische Quantifizierung der Zusammenhänge durch simulative Untersuchungen erfolgen. Das Netz wird mit Daten aus Experimenten befüllt. Im folgenden Abschnitt wird das Bayes'sche Netz für die Verkehrszustandsschätzung modelliert.

4.1.1 Knotenwahl und kausale Zusammenhänge

In [24] wird ein Bayes'sches Netz zur Lösung des Problems entwickelt. Im Fokus steht dabei die Fusion von Bluetooth- und C2X-Geschwindigkeitsmessungen zur Approximation der tatsächlichen mittleren Momentangeschwindigkeit. Im Folgenden soll diese zur Vereinfachung als tatsächliche Geschwindigkeit bezeichnet werden. Zur Verbesserung der Bluetooth-Detektion wird zudem die letzte Bluetooth-Detektion sowie die geschätzte Geschwindigkeitsdifferenz ins Bayes'sche Netz aufgenommen. Insgesamt entsteht das in Abbildung 3.6 gezeigte Netz.

An dieser Stelle soll kritisch hinterfragt werden, wieso das Bayes'sche Netz so gewählt wurde und ein alternativer Vorschlag für ein Bayes'sches Netz gemacht werden. Nachdem eine Vielzahl von möglichen Einflussfaktoren wie in 3.5 identifiziert wurden, sollen diese hier nochmals analysiert werden. Bei der näheren Auswahl der zu betrachtenden Parameter ist nicht nur zu klären, welche Größen relevant sind, sondern es muss auch geprüft werden, welche erfassbar und hinreichend genau messbar sind. Schließlich muss eine Wahl getroffen werden, welche Einflussfaktoren im Modell berücksichtigt werden

sollen. Die Faktoren, die die tatsächliche Geschwindigkeit v beeinflussen, sind in [24]:

- Verkehrsdichte D ,
- Lichtsignalanlage TLC ,
- Belegung O .

In dieser Arbeit sollen diese Größen nochmals genauer untersucht und die Modellierung des Bayes'schen Netzes analysiert werden. Verkehrsdichte und Belegung sind räumliche Parameter und stehen in direktem Zusammenhang. Auch die Verkehrsstärke Q steht in direkter Relation zu -dichte und Belegung. Problematisch ist, dass alle drei praktisch schwer erfassbar sind, sofern nicht auf umfangreichere Infrastruktureinrichtungen wie Kameras oder Induktionsschleifen zurückgegriffen werden soll. Zur Verwendung von Auslastung, Verkehrsstärke und -dichte muss jeweils eine Schätzung der mittleren räumlichen oder zeitlichen Fahrzeuganzahl möglich sein. Da die Infrastruktur lediglich mit einer C2X-fähigen Roadside Unit und einem Bluetooth-Empfänger ausgestattet ist, ist eine exakte Erfassung der Fahrzeuganzahl nicht möglich. Die Anzahl der C2X-fähigen und mit Bluetooth-Geräten ausgestatteten Fahrzeuge ist hingegen erfassbar. Diese machen jedoch nur einen Teil der Fahrzeuge aus. Während die zufällige Ausstattung mit Bluetooth und C2X beim Einsetzen nach einem festgelegten Ausstattungsgrad diesen zwar im Mittel erreicht, unterliegt das Verhältnis von Fahrzeugen mit und ohne Ausstattung zur Echtzeit hohen Schwankungen. Somit ist auch die Hochrechnung anhand der Ausstattungsgrade fehlerbehaftet. Ohne eine zuverlässige Schätzung der Fahrzeuganzahl lassen sich Verkehrsstärke und -dichte nicht berechnen. Sie werden daher nicht ins Bayes'sche Netz aufgenommen. Ihre Erfassbarkeit ist in der praktischen Anwendung schwer zu gewährleisten und soll auch in der simulativen Untersuchung nicht vorausgesetzt werden, um möglichst realistische Ergebnisse zu erhalten. Ähnliches gilt für die Belegung, die stets fehlerhaft ist. Es wird nämlich keine Belegung erkannt, wenn der Detektionsbereich von einem nicht ausgestatteten Fahrzeug belegt wird. So können keine zuverlässigen Aussagen gemacht werden.

Simulativ hingegen sind diese Größen sehr wohl erfassbar. Daher wurde die Verkehrsstärke Q testweise ins Bayes'sche Netz integriert. Erste Ergebnisse konnten jedoch keine signifikante Verbesserung der Schätzung verzeichnen. Da die Verkehrsstärke die Schätzung weder verbessern kann, noch den Test auf praktische Integrierbarkeit im Verkehr besteht, wird von der Aufnahme ins Bayes'sche Netz abgesehen.

Der Einfluss der Lichtsignalanlage auf die Geschwindigkeit hingegen ist praktisch messbar. Wird eine Roadside Unit zur Erfassung der C2X-Kommunikation eingeplant, die überdies mit Rechenkapazität zur Datenfusion ausgestattet ist, kann diese auch als mit der Lichtsignalanlage verknüpft angenommen werden. Es sei vorausgesetzt, dass Lichtsignalanlage, Bluetooth-Empfänger und Roadside Unit verknüpft sind und Informationsaustausch möglich ist. Damit ist der Zustand der Lichtsignalanlage zu jedem Zeitpunkt bekannt.

Es ist davon auszugehen, dass die Phase der Lichtsignalanlage die tatsächliche Geschwindigkeit beeinflusst. Zeigt die Lichtsignalanlage *ROT*, müssen die Fahrzeuge anhalten und verringern ihre Geschwindigkeit bzw. bleiben stehen. Während der *GRÜN*-Phase dürfen die Fahrzeuge weiterfahren und setzen ihre Fahrt über die Kreuzung fort. Es ist daher eine Abhängigkeit der Geschwindigkeit von der Phase der Lichtsignalanlage gegeben, sodass die Phase der Lichtsignalanlage ins Modell integriert werden soll. Diese Abhängigkeit wird in Abschnitt 4.1.2 genauer untersucht.

Auch die Einflussfaktoren, die zur Verbesserung der Bluetooth-Detektion V_{BT} identifiziert wurden, sollen einer genaueren Analyse unterzogen werden. Es sei angemerkt, dass die Notation in [24] von der hier verwendeten abweicht. So wird die Geschwindigkeit V_{V2X} im Folgenden als V_{C2X} bezeichnet. Die letzte Bluetooth-Detektion heißt in [24] V'_{BT} , soll hier aber als V_{pre} bezeichnet werden, um Missverständnisse mit der Ableitung zu vermeiden.

Zur Verbesserung der Bluetooth-Detektion herangezogen wurden:

- letzte Bluetooth-Detektion V_{pre} ,
- geschätzte Geschwindigkeitsdifferenz ΔV ,
- geschätzte Zeitlücke ΔT .

Naheliegender ist eine Betrachtung des vorausfahrenden Fahrzeugs anhand dessen Geschwindigkeit, Geschwindigkeitsdifferenz zum Folgefahrzeug oder der Zeitlücke zwischen beiden. Doch wie in [24] festgestellt sind die Informationen aus geschätzter Zeit- und Geschwindigkeitslücke redundant. Weiter lässt sich argumentieren, dass auch Geschwindigkeitsdifferenz und letzte Detektion redundante Ergebnisse liefern, da die geschätzte Geschwindigkeitslücke nichts anderes ist als die Differenz

$$\Delta V = V_{pre} - V_{BT}.$$

Diese Informationen werden jeweils aus den erfassten Ein- und Ausfahrtszeitpunkten berechnet und beruhen daher auf denselben Daten. Vor diesem Hintergrund soll auf die Geschwindigkeitsdifferenz verzichtet werden und lediglich die letzte Bluetooth-Detektion ins Bayes'sche Netz aufgenommen werden.

Die Vermutung, dass es hinreichend ist, die Geschwindigkeiten ins Netz aufzunehmen, hat sich in Experimenten mit unterschiedlichen Bayes'schen Netzen bestätigt. Wird einem Bayes'schen Netz, das die beiden Geschwindigkeiten enthält, noch die Geschwindigkeits- oder Zeitdifferenz hinzugefügt, so werden dadurch stets vergleichbare Ergebnisse erzielt.

Die zu beobachtenden Zufallsvariablen sind damit:

- Phase der Lichtsignalanlage,
- tatsächliche mittlere Momentangeschwindigkeit,
- detektierte Bluetooth-Geschwindigkeit,
- detektierte C2X-Geschwindigkeit,
- detektierte Bluetooth-Geschwindigkeit des vorausfahrenden Fahrzeugs.

Die kausalen Zusammenhänge zwischen den Knoten wurden in 3.5 thematisiert und werden nun nochmals wiederholt. Dafür werden die einzelnen Faktoren betrachtet und Ursachen und Wirkungen identifiziert.

- Die detektierten Geschwindigkeiten hängen jeweils von der tatsächlichen Geschwindigkeit ab, da es sich bei den Detektionen um einen Messvorgang handelt. Gleich, ob Bluetooth-Detektion oder C2X-Kommunikation zur Geschwindigkeitsschätzung genutzt wird: Die Schätzung ist stets eine mehr oder minder genaue Messung der tatsächlichen Geschwindigkeit.
- Die Phase der Lichtsignalanlage steuert, ob Fahrzeuge halten oder die Kreuzung queren. Daher nimmt sie direkt Einfluss auf die tatsächliche Geschwindigkeit im Detektionsbereich.
- Weniger offensichtlich sind die kausalen Beziehungen zu der detektierten Geschwindigkeit des vorausfahrenden Fahrzeugs. Während die tatsächliche Geschwindigkeit des vorausfahrenden Fahrzeugs die tatsächliche Geschwindigkeit des Folgefahrzeugs beeinflusst, so gilt dies nicht für die Detektionsgeschwindigkeit. Sie wird hingegen als Einflussfaktor auf die Bluetooth-Detektion angenommen, da sich die Detektionszeitpunkte eines Fahrzeugs und die des Folgefahrzeugs teilweise überschneiden können.

Es ist zu anzunehmen, dass die C2X-Geschwindigkeit eine bessere Schätzung der tatsächlichen Geschwindigkeit ist als die detektierte Bluetooth-Geschwindigkeit. Aus diesem Grund wird die detektierte Geschwindigkeit des vorausfahrenden Fahrzeugs nur im Falle einer Bluetooth-Detektion betrachtet.

Damit ergibt sich für das Bayes'sche Netz die Struktur wie in Abbildung 4.1 dargestellt.

Das Bayes'sche Netz soll es ermöglichen, die tatsächliche Geschwindigkeit anhand der Messgeschwindigkeiten V_{BT} und V_{C2X} sowie V_{pre} und TLC zu schätzen. Es soll daher die bedingte Wahrscheinlichkeit $P(V|V_{BT}, V_{C2X}, V_{pre}, TLC)$ bestimmt werden. Die Zufallsvariablen V_{BT} , V_{C2X} , V_{pre} und TLC sind dabei stets bekannt, wenn ein entsprechend ausgestattetes Fahrzeug im Detektionsbereich ist. Sie können als Evidenzen für die unbekannte tatsächliche Geschwindigkeit V genutzt werden.

Mit der Fusionsgleichung in Bayes'schen Netzen gilt für die Berechnung für alle $V = v$,

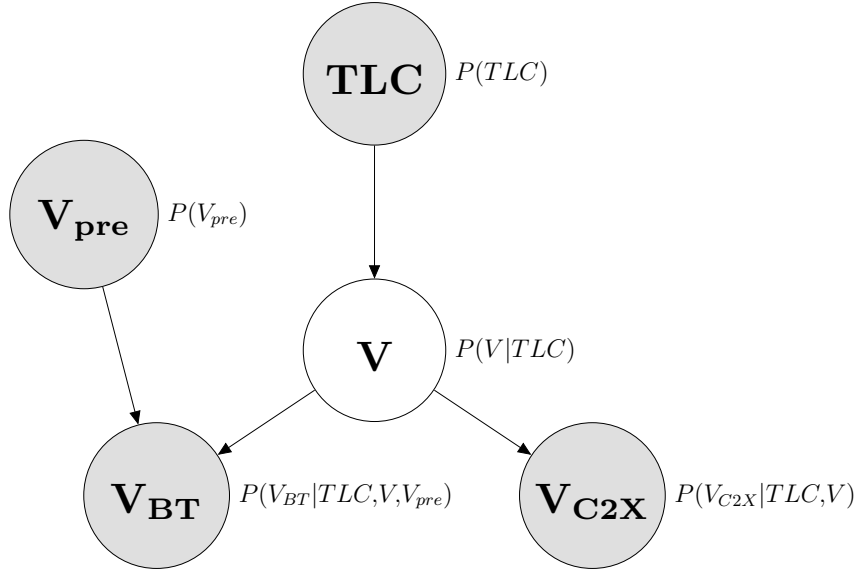


Abbildung 4.1: Struktur des modellierten Bayes'schen Netzes mit bedingten Wahrscheinlichkeiten für alle Knoten.

Bedeutung der Knoten: TLC (traffic light control): Phase der Lichtsignalanlage, V_{pre} : Geschwindigkeit des vorausfahrenden Fahrzeugs, V : tatsächliche mittlere Momentangeschwindigkeit, V_{BT} : Bluetooth-basierte Geschwindigkeitsschätzung, V_{C2X} : C2X-basierte Geschwindigkeitsschätzung.

$$V_{BT} = v_{BT}, V_{C2X} = v_{C2X}, V_{pre} = v_{pre} \text{ und } TLC = tlc$$

$$P(v|v_{BT}, v_{C2X}, v_{pre}, tlc) = \alpha \cdot P(v|tlc) \cdot P(v_{pre}) \cdot P(v_{BT}|v, v_{pre}, tlc) \cdot P(v_{C2X}|v, tlc). \quad (4.2)$$

Dabei ist α ein Normierungsfaktor, der so zu wählen ist, dass für die Menge $\{v_1, \dots, v_n\}$ aller Realisierungen von V sichergestellt ist:

$$\sum_{i=1}^n P(v_i|v_{BT}, v_{C2X}, v_{pre}, tlc) = 1. \quad (4.3)$$

Dabei setzt sich die Fusionsgleichung aus den A-priori-Wahrscheinlichkeiten $P(v|tlc)$ und $P(v_{pre})$ und den Sensorlikelihoods $P(v_{BT}|v, v_{pre}, tlc)$ sowie $P(v_{C2X}|v, tlc)$ zusammen. Eine Sensorlikelihood ist ein Maß dafür, wie genau ein Sensor die tatsächlichen Gegebenheiten abbilden kann. Für eine bestimmte tatsächliche Geschwindigkeit gibt sie an, welche Geschwindigkeit detektiert wird und mit welcher Wahrscheinlichkeit dies geschieht.

Um die tatsächliche Geschwindigkeit zu schätzen, können die beiden durch Bluetooth bzw. C2X detektierten Geschwindigkeiten verwendet werden. Sind zu einem Zeitpunkt

jedoch nur Daten einer Detektionsart verfügbar, also entweder Bluetooth oder C2X, kann die Fusionsgleichung nicht angewendet werden. Da die beiden Detektionsgeschwindigkeiten jedoch unabhängig voneinander sind, reicht eine von ihnen aus, um eine Schätzung der tatsächlichen Geschwindigkeit zu ermöglichen. Die Fusionsgleichung reduziert sich dann um die jeweils unbenötigten Faktoren $P(v_{C2X}|v, tlc)$ bzw. $P(v_{BT}|v, v_{pre}, tlc)$ und $P(v_{pre})$.

Ist nur eine Bluetooth-Detektion vorhanden, so gilt:

$$P(v|v_{BT}, v_{pre}, tlc) = \alpha_{BT} \cdot P(v|tlc) \cdot P(v_{pre}) \cdot P(v_{BT}|v, v_{pre}, tlc). \quad (4.4)$$

Analog gilt für den Fall, dass nur C2X-Daten detektiert werden:

$$P(v|v_{C2X}, tlc) = \alpha_{C2X} \cdot P(v|tlc) \cdot P(v_{C2X}|v, tlc). \quad (4.5)$$

Es müssen dabei α_{BT} und α_{C2X} so gewählt werden, dass gilt $\sum_{i=1}^n P(v_i|v_{BT}, v_{pre}, tlc) = 1$ bzw. $\sum_{i=1}^n P(v_i|v_{C2X}, tlc) = 1$.

Sind beide Geschwindigkeiten vorhanden, können die beiden Detektionen mit Gleichung 4.2 fusioniert werden. So können alle verfügbaren Informationen ausgenutzt werden.

Um das Bayes'sche Netz für solche Inferenzen zu nutzen, müssen zunächst die möglichen Wertebereiche für die Zufallsvariablen festgelegt werden. Anschließend können Wahrscheinlichkeitsverteilungen bestimmt werden.

4.1.2 Diskretisierung der Knoten

Ist die Struktur des Bayes'schen Netzes festgelegt, müssen die Knoten noch mit einem geeigneten Wertebereich versehen werden. Eine Diskretisierung der Wertebereiche bietet sich an, da so in Experimenten das Erfassen der Vorkommen einzelner diskreter Klassen möglich wird. So können Wahrscheinlichkeitsverteilungen durch Feststellen relativer Häufigkeiten bestimmt werden. Daher sollen Diskretisierungen für die Geschwindigkeiten V , V_{BT} , V_{C2X} und V_{pre} festgelegt werden. Auch für die Lichtsignalanlage TLC sollen diskrete Klassen definiert werden.

Um eine passende Diskretisierung für die Geschwindigkeiten zu finden, können zunächst allgemeine Feststellungen über deren Wertebereich gemacht werden. Da jeweils innerstädtische Szenarien untersucht werden sollen, liegt die Höchstgeschwindigkeit bei 50 km/h $\approx 13,89$ m/s. In SUMO sind jedoch auch Überschreitungen der Höchstgeschwindigkeiten möglich. Die Geschwindigkeit v soll daher im Bereich $0 \text{ m/s} \leq v \leq 20 \text{ m/s}$ abbildbar sein.

Diskrete Klassen können intuitiv gewählt werden, indem die Geschwindigkeiten auf die nächste ganze Zahl abgerundet werden. Mit 20 solcher Klassen sind folglich Geschwindigkeiten bis zu 20 m/s abgebildet, was etwa 72 km/h entspricht. Die Maximalgeschwindigkeit sollte dabei stets an das verwendete Szenario angepasst werden. Für eine maximale Geschwindigkeit von 50 km/h, etwa 13,89 m/s, sind 13 Klassen ausreichend. Die diskreti-

sierte maximale Geschwindigkeit ist dabei 13. Um auch Überschreitungen der Höchstgeschwindigkeit korrekt abzubilden, sollen auch Geschwindigkeiten bis $20 \text{ m/s} = 72 \text{ km/h}$ berücksichtigt werden. Dafür werden 20 Klassen benötigt. In Einzelfällen können auch noch höhere Geschwindigkeiten auftreten. Diese sollen der höchsten Klasse 19 zugeordnet werden. Folglich werden alle Geschwindigkeiten ab 19 m/s Klasse 19 zugeordnet. So werden die Geschwindigkeiten wie folgt in Klassen mit Klassenbreite 1 m/s eingeteilt:

Diskrete Klasse	gemessene Geschwindigkeit
0	$0 \text{ m/s} \leq v < 1 \text{ m/s}$
1	$1 \text{ m/s} \leq v < 2 \text{ m/s}$
...	...
19	$19 \text{ m/s} \leq v$

Tabelle 4.1: Zuordnung der Geschwindigkeiten in diskrete Klassen.

Für die Lichtsignalanlage bietet sich die Unterscheidung der Phasen *ROT* und *GRÜN* an. Bei *ROT* bremsen Fahrzeuge ab und kommen zum Halt, während bei *GRÜN* die Fahrzeuge beschleunigen und die Kreuzung dann queren. Stehen Fahrzeuge an der Kreuzung, so beschleunigen diese bis nach einigen Sekunden ein freier Verkehrsfluss entstehen kann. Diese Beobachtung motiviert, die Phase *GRÜN* nochmals in zwei verschiedene Klassen zu unterteilen: Eine Klasse für die ersten $7,5 \text{ s}$ der Grünphase, eine weitere für die restliche Phasendauer. Die Einteilung nach $7,5 \text{ s}$ ist nach experimentellen Untersuchungen ausgewählt. Die Phase *GELB* ist in SUMO so realisiert, dass Fahrzeuge passieren. Nur noch weit von der Kreuzung entfernte Fahrzeuge leiten den Abbremsvorgang ein. Da nur der Bereich nahe der Kreuzung betrachtet wird, wird im Folgenden die Phase *GELB* stets zur Phase *GRÜN* gezählt. *GELB* ist insgesamt eine Überlagerung von *GRÜN* und *ROT* und könnte ebenfalls detailliert analysiert werden. Dafür wäre auch zu klären, inwiefern sich das Fahrverhalten bei *GELB* in Realität und Simulation unterscheidet.

In Abbildung 4.2 sind beispielhaft zwei Geschwindigkeitsverläufe während der *GRÜN*-Phase dargestellt. Die Daten stammen aus Simulationsläufen des in Abschnitt 2.3.3 beschriebenen RiLSA-Szenarios. Sie zeigen, wie die Geschwindigkeit zunächst ansteigt und sich erst nach einigen Sekunden je nach Verkehrsfluss unterscheidet. Testläufe mit einer früheren oder späteren Unterteilung zeigen weniger signifikante Unterschiede in den Geschwindigkeitsverteilungen.

Um zu prüfen, wie die Phase der Lichtsignalanlage tatsächlich die Geschwindigkeitsverteilung beeinflusst, wurde experimentell bestimmt, wie sich die Geschwindigkeit nur in Abhängigkeit der Phase verhält. Im Graphen 4.3 sind die diskreten Wahrscheinlichkeitsfunktionen $P(v|tlc)$ für $tlc \in \{\text{rot}, \text{grün1}, \text{grün2}\}$ zu sehen.

Dabei ist auffällig, dass sich die Wahrscheinlichkeitsfunktionen je nach Phase deutlich unterscheiden. Die Geschwindigkeitswahrscheinlichkeitsfunktion während der *ROT*-Phase

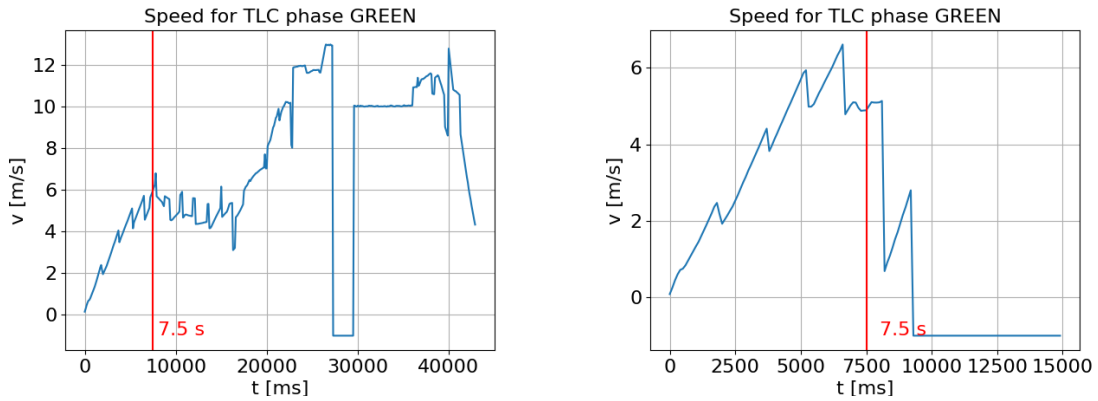


Abbildung 4.2: Verlauf der mittleren Momentangeschwindigkeit während der *GRÜN*-Phase. Simulationsergebnisse des RiLSA-Szenarios aus Abschnitt 2.3.3. Links: aus östlicher Richtung, rechts: aus südlicher Richtung.

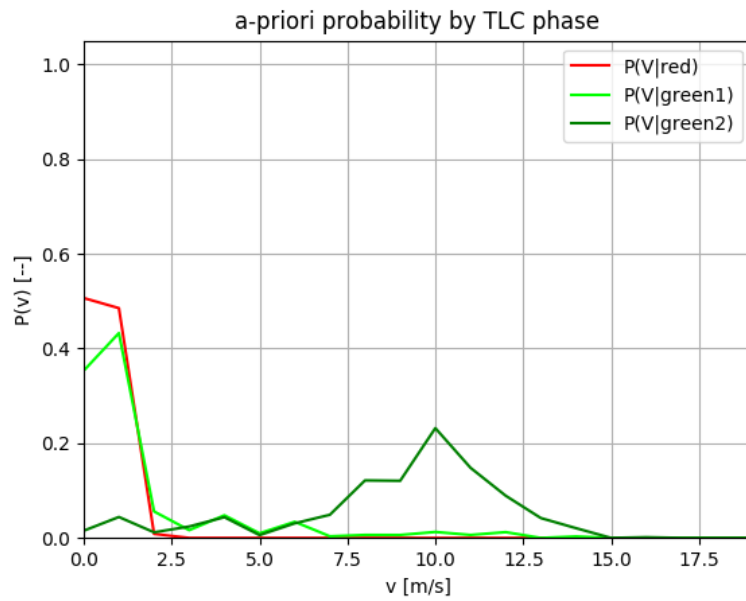


Abbildung 4.3: Empirische Wahrscheinlichkeitsfunktionen für die Geschwindigkeit in Abhängigkeit der Phase der Lichtsignalanlage. Rot: mittlere Momentangeschwindigkeit im Bluetooth-Detektionsbereich bei *ROT*, hellgrün: bei *GRÜN* bis 7,5 s, dunkelgrün: ab 7,5 s.

zeichnet sich durch niedrige Geschwindigkeiten aus. Der erste Teil der *GRÜN*-Phase erinnert noch stark daran, während der zweite Teil der Phase wesentlich höhere Geschwindigkeiten beinhaltet und das gesamte Geschwindigkeitsspektrum umfasst. Die

Diskretisierung ist so zu wählen, dass signifikante Unterschiede in der Geschwindigkeitsverteilung auch abgebildet werden können.

Auch die letzte Bluetooth-Detektion soll mit einer geeigneten Diskretisierung versehen werden. Es genügt, sie in zwei Klassen einzuteilen, da wie in Abbildung 4.4 dargestellt vor allem niedrige Geschwindigkeiten auftreten. Die Unterscheidung wird nur zwischen niedriger und hoher Geschwindigkeit getroffen. Die Klassen werden als *schnell* und *langsam* bezeichnet und deren Grenze bei 3 m/s definiert. Sei v_{det} die letzte detektierte Bluetooth-Geschwindigkeit und v_{pre} ihr diskretisiertes Pendant. Dann gilt

$$v_{pre} = \begin{cases} \text{langsam,} & 0 \leq v_{det} < 3, \\ \text{schnell,} & 3 \leq v_{det}. \end{cases} \quad (4.6)$$

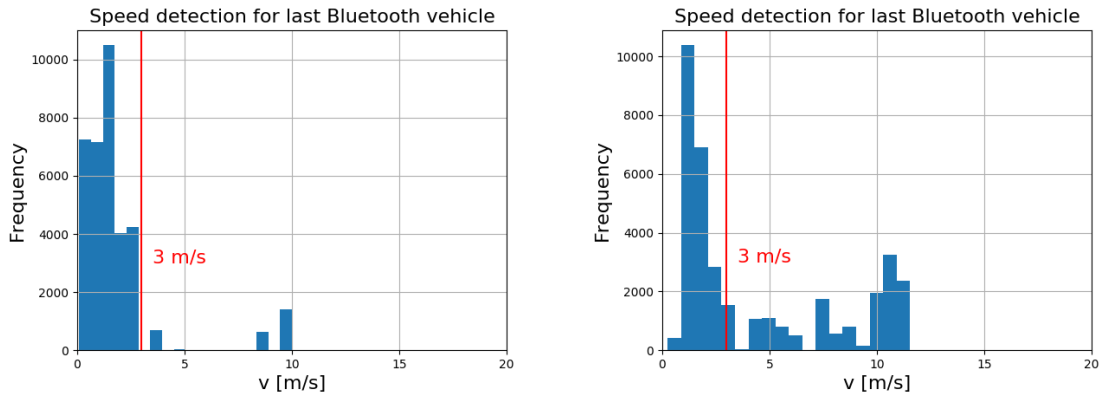


Abbildung 4.4: Histogramm für die jeweils letzten Bluetooth-Detektionen. Links: aus südlicher Richtung, rechts: aus östlicher Richtung.

Hier ist ein deutlicher Unterschied zur Unterteilung der Geschwindigkeiten von v , v_{BT} und v_{C2X} gegeben. Während für diese drei Geschwindigkeiten die diskreten Klassen von 0 bis 19 m/s verwendet werden, werden für v_{pre} nur die Klassen *schnell* und *langsam* unterschieden. Während v_{BT} stets geschätzt wird, wenn ein Fahrzeug detektiert wird, wird die letzte Bluetooth-Detektion v_{pre} gespeichert und so lange vorgehalten, bis ein weiteres Bluetooth-Fahrzeug den Bereich verlässt. Von jedem Fahrzeug wird nur eine einzige Schätzggeschwindigkeit angenommen. Daher ist der Datensatz für v_{pre} erheblich kleiner als der für v_{BT} .

Da weniger Daten für v_{pre} existieren, ist es schwieriger, die Wahrscheinlichkeitsverteilungen experimentell zu bestimmen. Da nur ausreichend quantifizierte Wahrscheinlichkeitsfunktionen gut zu verwenden sind, bietet es sich an, die gröbere Diskretisierung aus 4.6 zu verwenden.

Sind die diskreten Klassen bestimmt, kann die Fusionsgleichung 4.2 verwendet werden. Diese Gleichung wird angewandt, wenn sowohl Bluetooth- als auch C2X-Daten vorliegen. Liegen dagegen nur Daten von entweder C2X oder Bluetooth vor, ist ebenfalls eine Schätzung möglich. Dafür wird die Fusionsgleichung nur leicht modifiziert.

Für eine vorliegende C2X-Geschwindigkeit reicht es aus, die detektierte Geschwindigkeit mit allen möglichen Hypothesen über die tatsächliche Geschwindigkeit abzugleichen. Dazu wird die zur Phase der Lichtsignalanlage passende Wahrscheinlichkeitsverteilung $P(V|V_{C2X}, TLC)$ gesucht. Diese entspricht für $V = v$, $V_{C2X} = v_{C2X}$ und $TLC = tlc$

$$P(v|v_{C2X}, tlc) = \alpha_{C2X} \cdot P(v|tlc) \cdot P(v_{C2X}|v, tlc). \quad (4.7)$$

Der Faktor α_{C2X} stellt sicher, dass $\sum_{i=1}^n P(v_i|v_{C2X}, tlc) = 1$ gilt.

Ähnlich werden die Bluetooth-Wahrscheinlichkeiten bestimmt. Hier ist zusätzlich die detektierte Geschwindigkeit bei Ausfahrt des letzten Fahrzeugs zu berücksichtigen. Diese Geschwindigkeit wird gespeichert und bei jeder neuen Fahrzeugausfahrt aktualisiert. Somit ist sie in jedem Zeitschritt bekannt und zur Datenfusion verfügbar. Um mit Bluetooth-Daten die Geschwindigkeit zu berechnen, kann ausgenutzt werden, dass

$$P(v|v_{BT}, v_{pre}, tlc) = \alpha_{BT} \cdot P(v|tlc) \cdot P(v_{pre}) \cdot P(v_{BT}|v, v_{pre}, tlc). \quad (4.8)$$

Analog zum C2X-Fall ist α_{BT} ein Faktor zur Normierung der Wahrscheinlichkeit, sodass $\sum_{i=1}^n P(v_i|v_{BT}, v_{pre}, tlc) = 1$ gilt.

4.2 Implementierung

Das in 4.1 gemäß Abbildung 4.1 modellierte Bayes'sche Netz wird hier für die praktische Verwendung aufbereitet. Die Erklärungen folgen dabei der Reihenfolge der Teilschritte in der Simulation. Das Modell zur Verkehrszustandsschätzung gliedert sich dabei in drei Teile:

1. Zunächst werden in einer Verkehrssimulation Fahrzeuge detektiert und die zu erhebenden Kenngrößen gespeichert. Die Simulation muss mehrfach durchgeführt werden, um ausreichend Daten für den nächsten Schritt zu sammeln. Bei praktischer Anwendung bedeutet das, den Verkehr an der zu schätzenden Stelle mittels Bluetooth-Sensor und C2X-fähiger Roadside Unit sowie einer verlässlichen Messung der Geschwindigkeit aufzuzeichnen, bevor das Modell angewendet werden kann.
2. Im nächsten Schritt werden die gesammelten Daten ausgewertet und die Verbindungsmatrizen des Bayes'schen Netzes befüllt. Das Modell zur Datenfusion ist nun vorbereitet und quantifiziert, sodass es jetzt einsatzbereit ist.

3. Der letzte Schritt besteht darin, das Bayes'sche Netz in einer weiteren Simulation anzuwenden um eine Schätzung der mittleren Momentangeschwindigkeit zur Simulationszeit zu erhalten.

Die folgenden Abschnitte sind der Erklärung dieser drei Schritte gewidmet.

4.2.1 Detektion

Im ersten Schritt werden Simulationsläufe im Verkehrssimulator SUMO durchgeführt. Dafür werden die definierten Simulationsszenarien mehrfach ausgeführt. Als sinnvoll stellen sich ca. 500 Simulationsläufe heraus. Dabei werden Bluetooth- und C2X-Detektionen aufgezeichnet. Für C2X werden in jedem Zeitschritt Nachrichten versendet. Die Schrittweite der Simulationsschritte kann in SUMO zwischen 0,1 s und 1 s betragen. Für die betrachteten Szenarien wird eine Schrittweite von 0,1 s angenommen. Zur realistischeren Modellierung sollen auch Fehler beim Nachrichtenaustausch auftreten, die zur Nichtzustellung einer Nachricht führen. Es wird angenommen, dass nur ca. 90% der Nachrichten ankommen. Das ist mit einem Zufallszahlengenerator leicht realisierbar und eine pessimistische Annahme bei kleinen Reichweiten bis 30 m, wie Untersuchungen in [3] zeigen. Für niedrige Geschwindigkeiten bis 80 km/h sowie niedrige Reichweiten bis 50 m findet kaum Nachrichtenverlust statt. Die Empfangsrate der Nachrichten liegt über 95%. Da die Modellierung der erfolgreichen Nachrichtenübertragung eher pessimistisch ist, kann in der Realität diesbezüglich von besseren Ergebnissen als im Modell ausgegangen werden.

Bluetooth-Detektionen finden erheblich seltener nach dem in Abschnitt 2.2.2 beschriebenen Prinzip statt. Der Inquiry-Prozess wiederholt sich erst nach 2,56 s. Folglich kann ein Fahrzeug frühestens nach Ablauf dieser Zeit ein weiteres Mal detektiert werden.

Die Detektionen werden mittels eines Skripts gestartet, in dem die Szenarien definiert werden und schließlich mittels TraCI der Simulator SUMO gestartet wird. Dabei kann die Anzahl der durchzuführenden Simulationsläufe als Parameter übergeben werden. Dank TraCI ist es möglich, in jedem Zeitschritt die Fahrzeuge im Detektionsbereich zu untersuchen und bei gegebener Ausstattung alle relevanten Fahrzeugdaten festzuhalten. Die nur prozentual festgelegte Ausstattung mit C2X und Bluetooth wird für jedes neue Fahrzeug per Zufallsgenerator festgelegt. Auch die Nachrichtenübertragung von C2X-Nachrichten kann so durch eine gegebene Erfolgsquote modelliert werden.

Die Simulationsläufe dienen für beide Detektionsarten der Sammlung von Daten. Es werden für jedes Fahrzeug die echte und die mit Bluetooth bzw. C2X detektierte Geschwindigkeit aufgezeichnet. Weiter werden die Detektionszeiten, eine eindeutige ID zur Wiedererkennung des Fahrzeugs bzw. Bluetoothgerätes sowie der Zustand der Lichtsignalanlage zum Detektionszeitpunkt festgehalten. Letzterer wird gemäß den Diskretisierungen aus Abschnitt 4.1.2 angegeben.

Während der Simulationszeit werden fortlaufend Dateien erzeugt. Es wird für jeden Kreuzungsarm jeweils eine Datei mit Ergebnissen der beiden Detektoren - je ein Bluetooth-

und ein C2X-Detektor - erzeugt.

Der folgende Pseudocode stellt die Funktionsweise des Detektionsskripts dar.

Algorithmus 1 Detektion

```
1: for run < number_of_simulations do
2:   load_scenario()           ▷ load predefined scenario
3:   add_detectors()          ▷ add Bluetooth and C2X detectors
4:   traci.start()             ▷ start simulation
5:   while time < simulation_end do
6:     traci.simulationStep()   ▷ execute one simulation step at a time
7:     RegisterVehicles()       ▷ adds vehicles to the simulation randomly
                                ▷ equipped with Bluetooth / C2X
                                ▷ simulate Bluetooth and C2X detection
8:     BT_data = detectBluetoothVehicles(time)
9:     C2X_data = detectC2XVehicles(time)
                                ▷ separate files for C2X and Bluetooth detec-
                                ▷ tion data are written
10:    writeDetectedVehicleBluetoothToFile(run, BT_data)
11:    writeDetectedVehicleC2XToFile(run, C2X_data)
```

In den Detektionsdateien findet sich in jeder Zeile eine Detektion mit Zeitstempel, Fahrzeug- und Detektorinformationen. In Tabelle 4.2 sind die gespeicherten Informationen mit jeweils einem Beispiel für Bluetooth und C2X dargestellt. Dabei werden für C2X-Fahrzeuge zusätzlich verrauschte x- und y-Positionen gespeichert, die fehlerbehaftete Ortung simulieren.

Information	Bluetooth	C2X
Zeitstempel	71.5	76.6
Fahrzeug-ID	c197	c493
Fahrzeugnummer	3	4
Position x [m]	11448.292372	11444.405009
Position y [m]	13559.438813	13554.726585
verrauschte Position x [m]		11444.242375
verrauschte Position y [m]		13552.249291
Position auf Fahrbahn [m]	41.367510	46.520033
tatsächliche Geschwindigkeit [m/s]	3.935792	2.352885
detektierte Geschwindigkeit [m/s]	5	2.749815
Fahrbahn-ID	-51810854	-51810854
Fahrstreifen-ID	-51810854_3	-51810854_4
Detektor-ID	BT_em	C2X_em
Detektortyp	Bluetooth	C2X
Detektorposition auf Fahrbahn [m]	32	32
Detektorfahrbahn-ID	-51810854	-51810854
Detektionsreichweite [m]	30	200
Ausstattungsgrad [%]	30	10
Phase der Lichtsignalanlage	rot	grün seit $\leq 7,5$ s

Tabelle 4.2: Beispiel je einer gespeicherten Bluetooth- und C2X-Detektion im Szenario BS Forschungskreuzung.

4.2.2 Feststellen der relativen Häufigkeiten

Im nächsten Schritt werden aus den eben aufgezeichneten Daten relative Häufigkeiten bestimmt, mit denen das Bayes'sche Netz quantifiziert wird.

Zur Datenfusion wird die Formel aus Abschnitt 4.1.2 verwendet:

$$P(v|v_{BT}, v_{C2X}, v_{pre}, tlc) = \alpha \cdot P(v|tlc) \cdot P(v_{pre}) \cdot P(v_{BT}|v, v_{pre}, tlc) \cdot P(v_{C2X}|v, tlc).$$

Dabei ist α so gewählt, dass gilt:

$$\sum_{i=1}^n P(v_i|v_{BT}, v_{C2X}, v_{pre}, tlc) = 1.$$

Damit diese Fusionsgleichung zur Datenfusion verwendet werden kann, werden zunächst die A-priori-Wahrscheinlichkeiten $P(v|tlc)$ und $P(v_{pre})$ sowie die Sensorlikelihoods $P(v_{BT}|v, v_{pre}, tlc)$ und $P(v_{C2X}|v, tlc)$ bestimmt.

Diese Wahrscheinlichkeiten sind durch Zählen der relativen Häufigkeiten bestimmbar.

Dafür werden die Detektionen ausgewertet und ihren diskreten Klassen zugeordnet. In einem Skript werden die Ergebnisse der Detektionsläufe nacheinander eingelesen. Unter Verwendung der entsprechenden Diskretisierungen werden die Häufigkeiten des Auftretens der Ereignisse gezählt.

Die Werte der Lichtsignalanlage werden wie definiert eingelesen. Die Geschwindigkeiten werden gerundet, sodass sie je einer Klasse zugeordnet werden. Um die detektierte Geschwindigkeit des zuletzt ausfahrenden Fahrzeugs zu erhalten, lässt sich ausnutzen, dass die Einträge in den Detektionsdateien chronologisch geordnet sind. Daher reicht es, die detektierte Geschwindigkeit bei Ausfahrt eines Fahrzeugs zu speichern bzw. den gespeicherten Wert bei jeder Ausfahrt von Fahrzeugen aus dem Detektionsbereich zu aktualisieren. Dieser Wert wird gemäß der Diskretisierung noch in die Klassen *schnell* und *langsam* unterteilt. Wird die Anzahl des Auftretens der einzelnen Klassen durch die Gesamtzahl der Datenpunkte dividiert, ergibt sich die relative Häufigkeit der Ereignisse. Diese ist eine Näherung der Wahrscheinlichkeit für die jeweilige Kombinationen von Hypothese und Evidenz.

Sind die Häufigkeiten bestimmt, so werden die multidimensionalen Matrizen definiert und anschließend in XML-Dateien gespeichert. Eine Vektorschreibweise, bei der höchstens ein 1×20 -Vektor als XML-Element dargestellt und mit Attributen zur Einordnung in die diskreten Klassen versehen ist, wird genutzt. So sind besonders die mehrdimensionalen Matrizen gleichermaßen menschen- und maschinenlesbar.

Die Detektionsdateien werden nacheinander eingelesen und für jede Richtung wird eine Datei mit den Wahrscheinlichkeitsmatrizen gebildet. Im Skript für den Lernprozess werden dafür die eingelesenen Daten verwendet, um relative Häufigkeiten zu bestimmen.

Von entscheidender Bedeutung ist, wie diese bestimmt werden. Sie liefern schließlich die Wahrscheinlichkeitsverteilung für die Zufallsvariablen im Bayes'schen Netz. Im Skript finden sich verschiedene Methoden wie etwa `GET_VEHICLE_SPEED_DISTRIBUTION`, die zur Quantifizierung der einzelnen Wahrscheinlichkeiten genutzt werden. Jede Methode ist für das Bestimmen der relativen Häufigkeiten einer anderen Zufallsvariable zuständig.

Der Pseudocode 2 gibt ein Beispiel für eine solche Quantifizierung.

Die Verteilung der A-priori-Wahrscheinlichkeit des Knotens V wird hier in der Methode `GET_VEHICLE_SPEED_DISTRIBUTION` bestimmt. Dabei werden entsprechend der Diskretisierung Klassen gebildet. Dafür wird die tatsächliche Geschwindigkeit v auf einen ganzzahligen Wert abgeschnitten. Dieser entspricht der zugeordneten Klasse. Für die in jedem Zeitschritt auftretende Geschwindigkeit wird dies wiederholt und mit einem Counter die Anzahl der Vorkommen der Klassen sowie die Gesamtzahl der Datenpunkte gezählt. Schließlich genügt es für jede Klasse, die relative Häufigkeit als Quotienten aus Klassenhäufigkeit und Gesamtzahl der Geschwindigkeiten zu bestimmen.

Algorithmus 2 Bestimmen der relativen Häufigkeiten

```
1: function GET_VEHICLE_SPEED_DISTRIBUTION(vehicleData)
2:   for i,v in enumerate(vehicleData) do
3:     if  $v_{\min} \leq v \leq v_{\max}$  then  $\triangleright$  identify valid speeds
4:       v_class = int(v)  $\triangleright$  assign to discrete class
5:       v_temp[v_class] += 1  $\triangleright$  count occurrence of each discrete class ...
6:       cnt += 1  $\triangleright$  ... and overall number of speed data
7:   for i,v in enumerate(v_temp) do
8:     if cnt > 0 then
9:       v_hist[i] = float(v) / float(cnt)
                                    $\triangleright$  get relative frequency for each class from
                                   absolute frequencies
10:  return v_hist
```

4.2.3 Schätzen und Fusion der Daten

Ist das Bayes'sche Netz vollständig quantifiziert, kann es in einem erneuten Simulationslauf zur Datenfusion verwendet werden. Dazu wird die Simulation mit TraCI erneut gestartet. Es werden die Sensorlikelihood-XML-Dateien gelesen. Die Fahrzeug-Detektionen finden dann in jedem Simulationsschritt wie beim Schreiben der Detektionsdateien statt. Nun werden die Detektionen jedoch nicht in Dateien geschrieben, sondern direkt zu einer Geschwindigkeitsschätzung fusioniert. Die Datenfusion liefert dabei stets Ergebnisse, wenn wenigstens ein mit Bluetooth oder C2X ausgestattetes Fahrzeug vorhanden ist. Nur wenn weder C2X- noch Bluetoothdaten verfügbar sind, kann keine Geschwindigkeitsschätzung in Echtzeit durchgeführt werden.

Aus den C2X-Nachrichten und der Bluetooth-Detektion werden Evidenzen für die Schätzgeschwindigkeit erhalten. Der Zustand der Lichtsignalanlage wird erfasst. Er kontrolliert, welche Wahrscheinlichkeitsverteilung angenommen wird. Je nachdem, welche Ausstattungen mit Bluetooth oder C2X zum Simulationszeitpunkt vorhanden sind, wird die passende Fusionsgleichung aus Abschnitt 4.1.2 verwendet.

Sind die A-posteriori-Wahrscheinlichkeiten für alle Werte von v berechnet, wird eine Schätzung von v mit dem Maximum-A-posteriori-Schätzer (MAP) bestimmt und als Ergebnisgeschwindigkeit ausgewählt. Der MAP eignet sich hier, da er von allen diskreten Geschwindigkeitsklassen $v = 0, \dots, 19$ diejenige liefert, für die $P(v|v_{BT}, v_{C2X}, v_{pre}, tlc)$ maximiert wird. Damit wird die Wahrscheinlichkeit für das Eintreten der beobachteten Zustände $v_{BT}, v_{C2X}, v_{pre}, tlc$ maximiert.

Dazu genügt es, die A-posteriori-Wahrscheinlichkeiten mit den Fusionsgleichungen für C2X, Bluetooth und deren Kombination aus den Beobachtungen und bestimmten

Wahrscheinlichkeiten zu berechnen. Schließlich ist deren Maximum

$$\hat{v} = \arg \max_v P(v|v_{BT}, v_{C2X}, v_{pre}, tlc) \quad (4.9)$$

zu finden. Dieses Maximum wird als Schätzung der Geschwindigkeit verwendet.

Die Fusion ähnelt der Detektion strukturell stark. Statt die Detektionen in Dateien zu speichern, werden sie jedoch direkt verarbeitet. Zur Fusion kann das Detektionsskript wiederverwendet werden. Insgesamt ergibt sich folgendes Skript mit Fusions- bzw. Detektionsmodus als wählbarem Parameter.

Algorithmus 3 Gemeinsames Detektions- und Fusionsskript

```

1: for run < number_of_simulations do
2:   load_scenario()                                ▷ load predefined scenario
3:   add_detectors()                                ▷ add Bluetooth and C2X detectors
4:   traci.start()                                  ▷ start simulation
5:   while time < simulation_end do
6:     traci.simulationStep()                        ▷ execute one simulation step at a time
7:     RegisterVehicles()                            ▷ adds vehicles to the simulation randomly
                                                    equipped with Bluetooth / C2X
8:     BT_data = detectBluetoothVehicles(time)
9:     C2X_data = detectC2XVehicles(time)
                                                    ▷ simulate Bluetooth and C2X detection
10:    if mode == detection then
11:      writeDetectedVehicleBluetoothToFile(run, BT_data)
12:      writeDetectedVehicleC2XToFile(run, C2X_data)
                                                    ▷ separate files for C2X and Bluetooth detec-
                                                    tion data are written
13:    else if mode == fusion then
14:      fused_data = fuseData(BT_data, C2X_data)
                                                    ▷ use Bayesian net to execute data fusion
15:      writeFusionResultsToFile(fused_data)
16: if mode == fusion then
17:   evaluate()

```

Die Simulation wird mit TraCI gestartet. In jedem Zeitschritt werden die Detektoren ausgelesen und Bluetooth-Detektionen erfasst sowie C2X-Nachrichten ausgelesen. Diese Informationen werden jedoch nicht gespeichert, sondern direkt für eine Geschwindigkeits-schätzung verwendet. In Fusionsdateien werden für jeden Zeitschritt die Ergebnisse der Fusion gespeichert. Weitere Daten wie die geschätzten Geschwindigkeiten, der Zustand der Lichtsignalanlage und die letzte Bluetooth-Detektion werden ebenfalls festgehalten. Auch eine Berechnung der mittleren Momentangeschwindigkeit im relevanten Detektions-

bereich wird durchgeführt, sodass die Geschwindigkeitsschätzung mit den tatsächlichen Geschwindigkeiten verglichen werden kann. So entsteht ein Rahmen zur Analyse des Modells, mit dem die Ergebnisse später ausgewertet werden können.

Mit der Implementierung des Modells kann dieses nun auf den vorbereiteten Szenarien getestet werden. Die simulativen Untersuchungen folgen im nächsten Kapitel 5.

5 Simulative Untersuchungen

In diesem Kapitel kommt das soeben vorgestellte Modell zum Einsatz. Dabei wird erklärt, wie die Szenarien in SUMO ausgeführt werden und wie die Implementierung des Modells verwendet werden kann. In diesem Kontext wird die Modellierung der verwendeten Detektoren, die Verknüpfung von Modell und Simulation sowie das Framework zur Auswertung der Ergebnisse erklärt. Schließlich werden die Simulationsläufe für alle drei untersuchten Szenarien ausgewertet und deren Ergebnisse analysiert.

5.1 Aufbau der Untersuchungen

Im Verkehrssimulator SUMO können die in Abschnitt 2.3.3 beschriebenen Szenarien simuliert werden. Sie definieren bereits das Streckennetz und die Routen der Fahrzeuge. Auch die Anzahl der Fahrzeuge mit den jeweiligen Routen ist durch die Szenarien gegeben. Es wurden dabei drei Szenarien definiert:

1. **Szenario Minimal:** Das Szenario besteht aus einer einzigen Straße, die geradeaus von Osten nach Westen verläuft. Sie ist frei von Knotenpunkten und Kurven und ermöglicht somit eine Untersuchung des Modells frei von geometrischen Beeinflussungen. Der Verkehrsfluss ist dabei so gewählt, dass Verkehrsstärke und Geschwindigkeit schrittweise das gesamte Spektrum von hoch bis niedrig durchlaufen. Die Geschwindigkeit reicht von 2 m/s bis 20 m/s, die Verkehrsstärke von 360 bis ca. 1500 Fahrzeugen pro Stunde. Da die Geometrie des Netzes stark idealisiert ist und sich der Verkehrsfluss in der Realität fließend und nicht abrupt ändert, ist diese Modellierung wenig realistisch. Sie bietet jedoch den Vorteil, dass das Modell so in Abhängigkeit von Geschwindigkeit und Verkehrsstärke getestet werden kann. Schwankungen aufgrund fließender Übergänge und geometrischer Gegebenheiten können so weitgehend vermieden werden. Da das Szenario keinen Knotenpunkt enthält, nimmt es eine Sonderrolle ein: Es existiert keine Lichtsignalanlage im Szenario und somit keine Phase, die ausgelesen und ins Modell mit aufgenommen werden kann. Das verwendete Bayes'sche Netz kann also als um den Knoten *TLC* reduziert angenommen werden oder dieser kann mit konstantem beobachteten Wert angenommen werden.
2. **Szenario RiLSA:** In diesem Szenario steht ein Knotenpunkt mit vier Kreuzungsarmen im Mittelpunkt. Diese führen symmetrisch nach Norden, Osten, Süden und

Westen. Wieder sind die Arme als gerade Strecken modelliert, um den geometrischen Einfluss gering zu halten. Der Verkehrsfluss unterscheidet sich dabei je nach Richtung, aus der die Fahrzeuge kommen. Während die Ost-West-Achse mit je Richtung 1000 Fahrzeugen pro Stunde stärker frequentiert ist, kommen aus Norden und Süden je 300 Fahrzeuge pro Stunde. In diesem Szenario wird der Verkehrsfluss über die gesamte Simulationszeit konstant gehalten.

3. **Szenario BS Forschungskreuzung:** Das dritte Szenario ermöglicht eine realitätsnahe Untersuchung des Modells. Das Straßennetz ist einer real existierenden Forschungskreuzung in Braunschweig nachempfunden. Dabei wurden die Geometrien exakt übertragen. Auch die Verkehrsnachfrage ist aus einer realen Messung übernommen. Sie entspricht etwa 360 Fahrzeugen aus dem Norden, 650 aus dem Süden sowie 580 aus östlicher und 470 Fahrzeugen aus westlicher Richtung.

Mit den Szenarien sind die Rahmenbedingungen für die simulative Untersuchung gegeben. Sie müssen noch mit Detektoren ausgestattet werden, um Bluetooth- und C2X-Detektionen zu ermöglichen. Dafür werden räumliche Detektoren, sogenannte *Edge-Detektoren*, in SUMO genutzt. In Abbildung 5.1 ist ein solcher Detektor dargestellt.

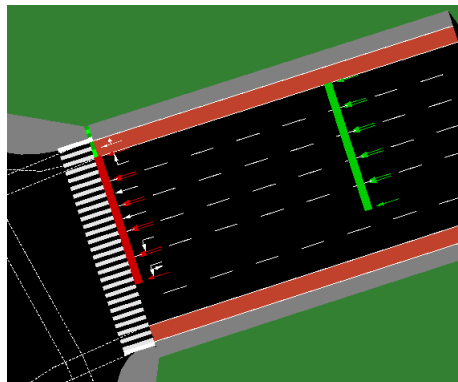


Abbildung 5.1: Edge-Detektor in SUMO, der zur Modellierung der C2X- und Bluetooth-Detektion verwendet werden kann. Einfahrtlinie (grün) und Ausfahrtlinie (rot) begrenzen den Detektionsbereich.

Die Detektoren werden modelliert, indem auf einer Kante ein Detektor durch die zugehörigen Fahrspuren mit Detektionsbeginn und -ende festgelegt wird. Sie lassen sich gleichermaßen für Bluetooth- und C2X-Detektion nutzen, sind aber je nach Anwendung anders zu parametrieren. Für die Bluetooth-Detektion ist die Reichweite klein zu wählen und stellt den gesamten Detektionsbereich des Bluetooth-Empfängers dar. Hier ist eine Reichweite von 30 m gewählt, was einem Klasse 2 Bluetooth-Gerät entspricht. Für die C2X-Kommunikation soll der Detektor als Roadside Unit (RSU) an der Kreuzung interpretiert werden. Die Reichweite kann höher gewählt werden. Hier wird eine Reichweite von 200 m gesetzt. Da sich die Edge-Detektoren nicht auf Kreuzungen einsetzen lassen,

werden zur Modellierung einer RSU vier Detektoren benötigt. Je einer wird auf jedem Kreuzungsarm integriert, sodass Detektionen in alle Himmelsrichtungen stattfinden können. Die gesammelten Informationen stellen dann den Kommunikationsbereich der RSU dar. Insgesamt werden acht solche Detektoren benötigt: Auf jedem Kreuzungsarm befindet sich je ein Bluetooth- und ein C2X-Detektor. Die beiden Detektoren sind so ausgerichtet, dass das Detektionsende jeweils etwa 1 m vor der Kreuzung auf genau dem selben Straßenquerschnitt liegt. Dies gilt für alle vier Himmelsrichtungen. Im Beispiel auf gerader Strecke sind die Detektoren ebenfalls so ausgerichtet, dass Bluetooth- und C2X-Detektionsbereich zeitgleich verlassen werden.

In einer separaten XML-Datei können diese Detektoren definiert und in die Simulation eingepflegt werden. Mit TraCI lässt sich in jedem Simulationsschritt auf die Detektoren zugreifen und auslesen, welche Fahrzeuge in deren Detektionsbereich sind. Für die detektierten Fahrzeuge wird dann geprüft, ob C2X- bzw. Bluetooth-Ausstattung vorliegt. Falls ja, können die Berechnungen zur Geschwindigkeitsschätzung wie erklärt durchgeführt werden.

Mit den simulierten Detektoren können in jedem Zeitschritt alle im Detektionsbereich befindlichen Fahrzeuge beobachtet werden. Tatsächlich können jedoch nicht alle Bluetooth-Geräte sofort detektiert werden [22]. In [24] wurde ein probabilistischer Ansatz zur Modellierung verwendet, bei dem angenommen wird, dass in 2,2 Sekunden etwa 70% der Bluetooth-Geräte detektiert werden. Diese Annahme wird hier übernommen und in den einzelnen Simulationsschritten als Bernoulli-Verteilung realisiert, sodass in jedem Zeitschritt die gleiche Detektionswahrscheinlichkeit angenommen wird. Bei einer Schrittweite von 0,1 s werden also 22 Simulationsschritte benötigt, um 70% der Fahrzeuge zu detektieren. Es gilt für die Detektionswahrscheinlichkeit p in jedem Schritt

$$\sum_{i=1}^{22} p(1-p)^{i-1} = 0,7.$$

Folglich ist $p \approx 0,0533$.

Für die C2X-Detektion wird eine Detektionswahrscheinlichkeit von 90% in jedem Zeitschritt angenommen. Untersuchungen in [3] zeigen, dass diese Annahme pessimistisch ist.

Damit die Bluetooth- und C2X-Detektionen besser nachvollziehbar sind, werden die Fahrzeuge in der Simulation entsprechend ihrer Ausstattung eingefärbt: Mit Bluetooth ausgestattete Fahrzeuge sind blau, mit C2X ausgestattete Fahrzeuge rot. Ist ein Fahrzeug mit beiden Technologien ausgestattet, ist es pink. Fahrzeuge ohne Ausstattung sind gelb. In Abbildung 5.2 ist ein Screenshot mit Fahrzeugen mit unterschiedlicher Ausstattung zu sehen.

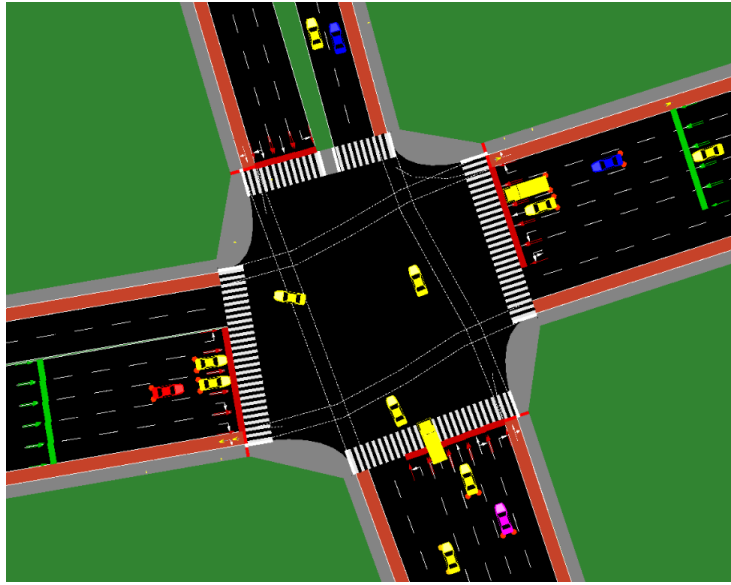


Abbildung 5.2: Darstellung von Fahrzeugen mit unterschiedlicher Ausstattung: Blau: Bluetooth, rot: C2X, pink: Bluetooth und C2X, gelb: ohne Ausstattung.

Sind Szenarien und Detektoren definiert, kann das Modell mit SUMO verknüpft und gestartet werden. Hierzu wird das in Abschnitt 4.2.1 beschriebene Skript verwendet. Darin werden zunächst Szenarien und Detektoren vorbereitet und dann mittels TraCI die Simulation gestartet. In jedem Simulationsschritt können nun kontrollierte Detektionen simuliert und in Dateien geschrieben werden. Die Simulation muss oft genug durchgeführt werden, um ausreichend Daten für den folgenden Lernalgorithmus zu sammeln. Etwa 800 Simulationsläufe haben sich als ausreichend erwiesen.

Im nächsten Schritt wird das Skript aus 4.2.2 genutzt, um aus den Detektionsdaten Wahrscheinlichkeitsverteilungen des Bayes'schen Netzes aus Kapitel 4 zu berechnen. Dafür werden die Daten ausgewertet und relative Häufigkeiten für die A-priori- und bedingten Wahrscheinlichkeiten festgestellt.

Schließlich kann mit Skript 4.2.3 die Simulation erneut ausgeführt werden. Wieder werden Detektionen durchgeführt. In diesem Schritt werden die Detektionen jedoch nicht aufgezeichnet, sondern direkt zur Simulationszeit zu einer Geschwindigkeitsschätzung fusioniert. Dabei werden Dateien erzeugt, in denen die geschätzten Geschwindigkeiten und die tatsächlichen Geschwindigkeiten in jedem Zeitschritt gespeichert werden. So kann nach Abschluss eines jeden Simulationslaufs direkt eine Evaluierung der Ergebnisse erfolgen.

5.2 Ergebnisse

In diesem Abschnitt werden die Ergebnisse der Simulation zusammengetragen und ausgewertet. Dabei werden Vollständigkeit und Genauigkeit analysiert und genauer untersucht, wie Größen wie Geschwindigkeit und Verkehrsstärke die Schätzungen beeinflussen. Auch die Bedeutung des C2X- und Bluetooth-Ausstattungsgrades wird hier erörtert.

Die Ergebnisse der Datenfusion lassen sich auf unterschiedliche Art und Weise bewerten. Zur Beurteilung von Datenqualität können nach [15] Genauigkeit und Vollständigkeit herangezogen werden.

1. Genauigkeit: Hiermit wird beschrieben, wie nah ein Schätzwert \hat{x} am korrekten Wert x liegt. Ist $\hat{x} = x$, wird von exakten Daten gesprochen.
2. Vollständigkeit: Die Vollständigkeit ist ein Maß dafür, ob die Datensammlung Einträge für alle Entitäten hat, die mit Daten versehen werden können. So können etwa fehlende Werte festgestellt werden. Ob unvollständige Datensätze zulässig sind und wie mit ihnen verfahren wird, unterscheidet sich je nach verwendetem Modell.

Zur Feststellung der Genauigkeit können unterschiedliche Metriken verwendet werden. Typisch sind der mittlere absolute Fehler (MAE) und der Root Mean Square Error (RMSE). Für $n \in \mathbb{N}$ Fehler e_1, \dots, e_n gilt dabei nach [6]:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |e_i|,$$
$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n |e_i|^2}.$$

MAE und RMSE unterscheiden sich darin, dass der RMSE größere Fehler stärker gewichtet, als der MAE. Da hier hohe und niedrige Abweichungen als Schätzung zu jeweils einem Zeitpunkt gleich gewichtet sind, soll bei detaillierten Betrachtungen der MAE verwendet werden.

5.2.1 Szenario Minimal

Zunächst wird im minimalistischen Szenario die grundlegende Funktionalität des Modells untersucht. Im Szenario existiert lediglich eine gradeaus führende Straße, an der Detektoren angebracht sind. Nach etwa 800 Detektionsläufen wurde das Lernverfahren angewendet, um die Wahrscheinlichkeiten zu bestimmen. Die relativen Häufigkeiten werden berechnet und in Matrizen gespeichert. Für $P(V_{C2X}|V)$ wird die Matrix M_{C2X} durch $(M_{C2X})_{i+1,j+1} = P(V_{C2X} = i|V = j)$ für alle $i, j = 0, \dots, 19$ bestimmt. Dabei hat M_{C2X} die Form

$$M_{C2X} = \left(\begin{array}{cccccccc|c} 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 & \\ 0 & 0.58 & \ddots & & & & & & & \vdots \\ 0 & 0.42 & 0 & \ddots & & & & & & \vdots \\ \vdots & \ddots & 0 & 0.6 & \ddots & & & & & \vdots \\ \vdots & & \ddots & 0.4 & 0 & \ddots & & & & \vdots \\ \vdots & & & \ddots & 0 & 0.57 & \ddots & & 0 & \\ \vdots & & & & \ddots & 0.43 & 0 & \ddots & 0 & \\ \vdots & & & & & \ddots & \ddots & \ddots & 0 & \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 0 & 0 & \\ \hline & & & & & & & & & 0 \end{array} \right). \quad (5.1)$$

Dabei ist $M_{i,j} = 0$ für alle $i \geq 15$ oder $j \geq 15$. Das ist dadurch zu erklären, dass Geschwindigkeiten über 14 m/s weder aufgetreten sind, noch detektiert wurden.

Weiter ist auffällig, dass nur in jeder zweiten Spalte Werte ungleich null auftreten. Das Szenario ist so modelliert, dass Fahrzeuge mit konstanter Geschwindigkeit fahren, die schrittweise erhöht wird. Die Schrittweite von 2 m/s bedingt, dass nur jede zweite Geschwindigkeitsklasse erreicht wird. Die einzelnen Spalten beschreiben das Auftreten bestimmter tatsächlicher Geschwindigkeiten, die jeweiligen Einträge die Detektionswahrscheinlichkeit in die jeweiligen Geschwindigkeitsklassen. Es ist zu sehen, dass jeweils nur zwei verschiedene Klassen detektiert werden: Die korrekte Klasse selbst und die nächsthöhere Geschwindigkeitsklasse. Dies spricht für eine recht genaue Detektion. Wie genau die Detektionen sind und welche Werte M_{C2X} annimmt, hängt davon ab, wie die Positionsbestimmung mittels GPS modelliert wird. Hier wurden Abweichungen von bis zu 3 m zufällig generiert. Größere Abweichungen führen zu einer größeren Streuung der Detektionen.

Der Großteil der Einträge der Matrix sind 0, sodass eine schwachbesetzte Matrix vor-

liegt. Sind in einer Spalte Einträge ungleich null vorhanden, ist deren Summe stets auf 1 normiert. Das Streichen der Spalten, die nur aus Nullen bestehen, würde somit eine stochastische Matrix erzeugen. Diese ist gerade durch die Eigenschaft definiert, dass die Spalten- oder Zeilensumme stets 1 ergibt.

Für die Bluetooth-Detektion entsteht eine Matrix mit ähnlichen Eigenschaften. Doch während für die C2X-Detektion in jeder Spalte maximal zwei positive Einträge existieren, sind es bei Bluetooth-Detektion mehr Einträge, deren Summe schließlich wieder 1 beträgt. So ist die erste Spalte mit positiven Einträgen auf zwei Nachkommastellen gerundet gerade

$$M_{BT_2} = (0 \quad 0.29 \quad 0.6 \quad 0.08 \quad 0.2 \quad 0.1 \quad 0 \quad \dots \quad 0)^T.$$

Die Abbildungen 5.3 und 5.4 zeigen die Sensorlikelihoods für Bluetooth- und C2X-Detektion.

In der Matrix ist dabei dargestellt, mit welcher Wahrscheinlichkeit die tatsächliche Geschwindigkeit v als Detektionsgeschwindigkeit v_{BT} bzw. v_{C2X} gemessen wird. Das Szenario Minimal nimmt hierbei eine Sonderrolle ein. Die Verkehrsflüsse sind so konzipiert, dass nur konstante Geschwindigkeiten vorliegen. Da sie Stück für Stück um 2 m/s verändert werden, liegen nur Geschwindigkeiten in jeder zweiten diskreten Klasse vor. Dadurch existieren nur Likelihoods für jede zweite Geschwindigkeitsklasse. Die farbig markierten Einträge zeigen die Wahrscheinlichkeitswerte für jede Kombination aus Detektions- und tatsächlicher Geschwindigkeit. Es fällt auf, dass bei der Bluetooth-Likelihood Geschwindigkeiten oft überschätzt werden. Die Schwankungen finden ausschließlich in diese Richtung statt und sind zudem oft wahrscheinlicher als korrektes Detektieren der Geschwindigkeit. Zudem zeichnet sich eine Häufung der Geschwindigkeiten bei Klasse 11 ab. Dies entspricht den in Abschnitt 2.2.2 festgestellten Eigenschaften der Bluetooth-Detektion, dass

- Detektionsgeschwindigkeiten größer oder gleich der tatsächlichen Geschwindigkeit sind,
- das absolute Maximum durch den 2,56 s langen Inquiry-Prozess mit Klasse 11 erreicht ist.

In der Matrix sind jedoch für eine tatsächliche Geschwindigkeit von 13 m/s auch Bluetooth-Detektionen von 13 m/s erfasst. Diese Schätzung wird gesetzt, wenn ein Fahrzeug mit Bluetooth detektiert wurde, aber keine zweite Erfassung stattfindet, bevor das Fahrzeug den Detektionsbereich verlässt. Da dies besonders bei schnellen Fahrzeugen möglich ist, wird in diesem Fall die Richtgeschwindigkeit von 50 km/h $\approx 13,89$ m/s angenommen.

Die C2X-Sensorlikelihood zeigt, dass die Geschwindigkeitsdetektion hier wesentlich

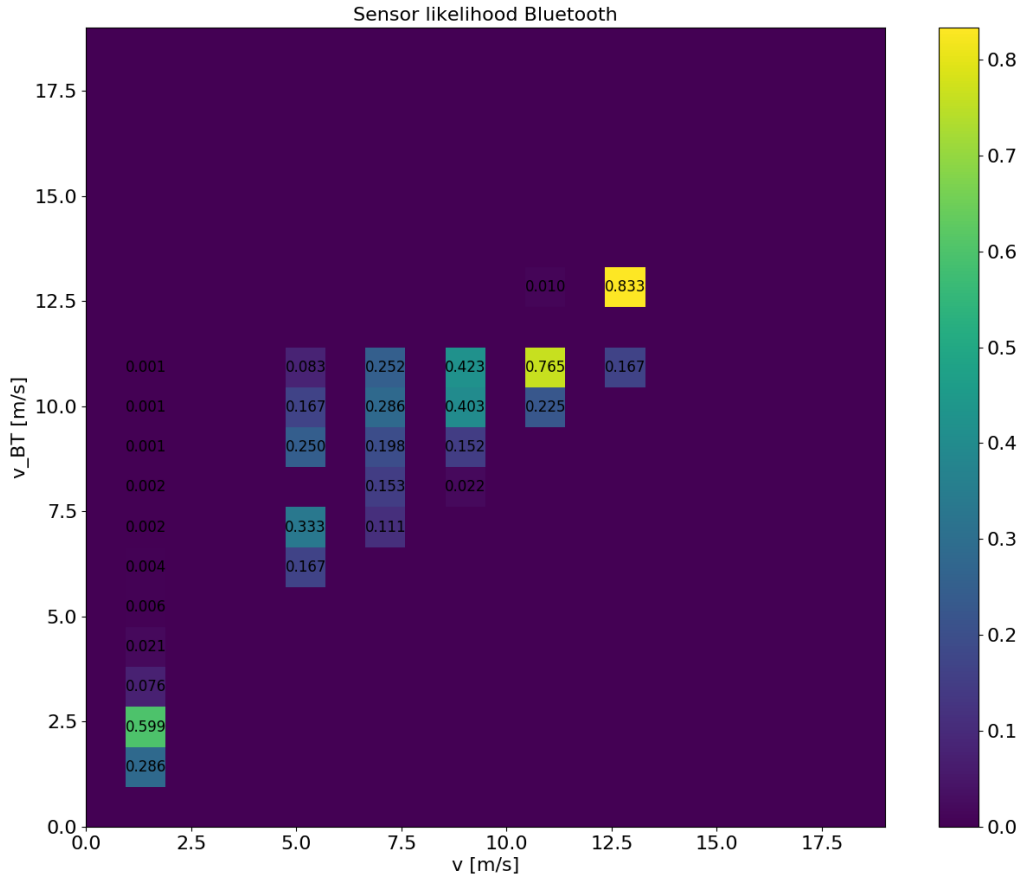


Abbildung 5.3: Sensorlikelihood für Bluetooth-Detektion.

besser funktioniert. Die Likelihoods erreichen alle bei der korrekten Detektionsgeschwindigkeit ihr Maximum mit über 50%. Auch die Streuung der Detektionen ist weniger groß als für die Bluetooth-Detektion. Es liegen jeweils noch Schätzungen der nächsthöheren Geschwindigkeitsklasse vor, Abweichungen um zwei oder mehr Geschwindigkeitsklassen gibt es jedoch nicht. Für eine andere Modellierung der GPS-Positionsbestimmung würden hier andere Ergebnisse auftreten.

Wurde die Datenfusion durchgeführt, lassen sich schließlich Genauigkeit und Vollständigkeit der Daten untersuchen. Für variierte C2X-Ausstattungsgrade bei fester Bluetooth-Penetrationsrate von 30% ergeben sich MAE, RMSE und Vollständigkeit wie

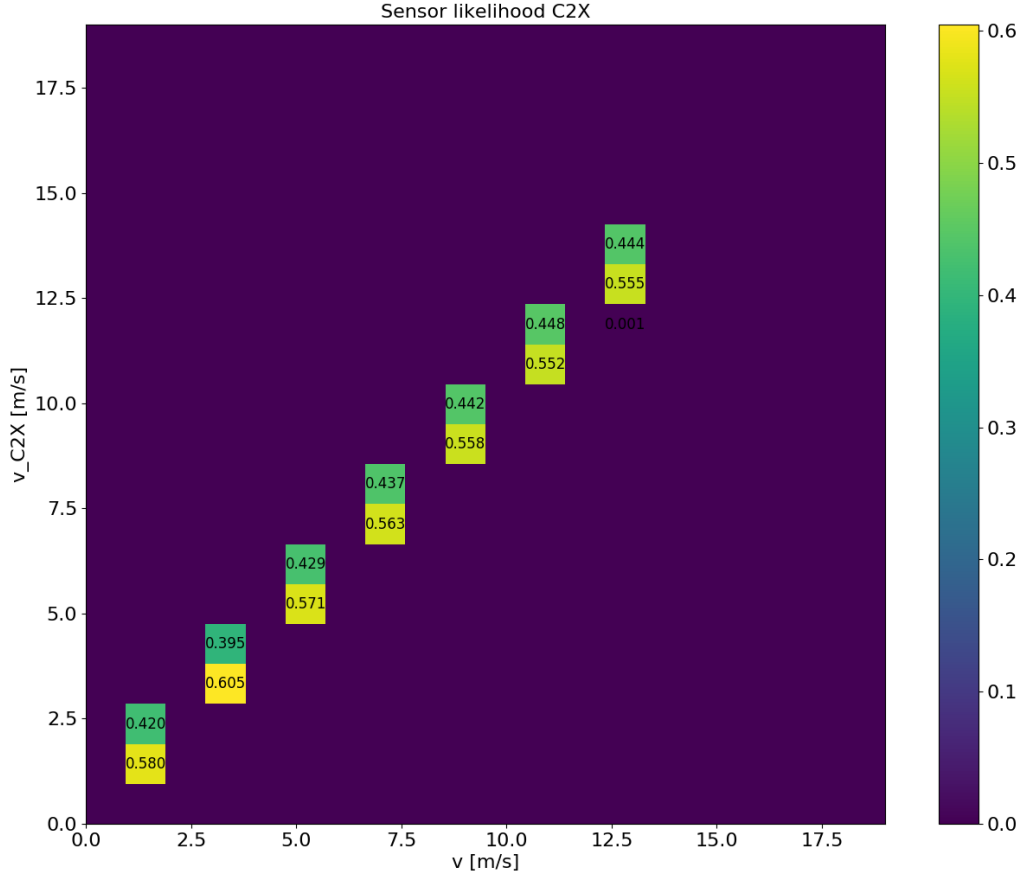


Abbildung 5.4: Sensorlikelihood für C2X-Detektion.

in Tabelle 5.1.

Die C2X-Penetrationsrate wird stückweise von 0 auf 100% erhöht. Da für die praktische Anwendung vor allem der Fall interessant ist, dass wenige C2X-fähige Fahrzeuge verfügbar sind, geschieht dies nicht linear, sondern mit starkem Fokus auf niedrige Ausstattungsgrade.

Die Genauigkeit ist ein Maß dafür, wie weit Schätzung und tatsächlicher Wert voneinander abweichen. Erhöht sich der Ausstattungsgrad, wird auch die Genauigkeit und Vollständigkeit der Ergebnisse erhöht. Ist die Genauigkeit mit $RMSE > 5$ m/s zunächst schlecht, fällt sie ab 10% Ausstattung unter 5 m/s, ab 20% unter 4 m/s und verbessert sich bei noch höherer Penetrationsrate weiter.

C2X-Ausstattung [%]	MAE [m/s]	RMSE [m/s]	Vollständigkeit [%]
0	5,27	6,65	48,8
1	4,94	6,38	50,58
5	4,1	5,67	55,99
10	3,35	4,96	61,54
20	2,37	3,87	72,19
50	1,28	2,02	89,71
100	0,98	0,98	99,71

Tabelle 5.1: Ergebnisse in Abhängigkeit der C2X-Ausstattungsgrade bei fester Bluetooth-Ausstattung von 30%.

Die Vollständigkeit liefert Aussagen darüber, wie oft das Modell eine Geschwindigkeitsschätzung aufstellen kann. Dies ist möglich, wenn entweder Bluetooth- oder C2X-ausgestattete Fahrzeuge vorliegen. Daher hängt sie direkt vom Ausstattungsgrad ab. Die Referenz hierfür sind alle Zeitpunkte, zu denen Fahrzeuge im Detektionsbereich sind. Dann liegen stets tatsächliche Geschwindigkeiten vor. Die Vollständigkeit gibt an, wie oft auch Schätzergebnisse für diese Werte existieren.

Durch die Bluetooth-Penetration kann selbst ohne C2X-fähige Fahrzeuge schon eine Vollständigkeit von fast 50% erreicht werden. Während der Ausstattungsgrad den Anteil der Fahrzeuge mit Bluetooth-Geräten beschreibt, bezeichnet die Vollständigkeit den Anteil der Zeitpunkte, zu denen Fusionsergebnisse vorliegen. Ein ausgestattetes Fahrzeug ist während der gesamten Zeitspanne detektierbar, die es zum Durchfahren des Detektionsbereichs benötigt. Somit kann ein Ausstattungsgrad von 30% zu einer deutlich höheren Vollständigkeit führen.

Ähnliches gilt für den C2X-Ausstattungsgrad von 100%. Es sind zwar alle Fahrzeuge C2X-fähig, die Modellierung der Kommunikation sieht jedoch vor, dass pro Zeitschritt nur 90% der Nachrichten zugestellt werden. Daher erreicht die Datenfusion keine Vollständigkeit von 100%.

Die Vollständigkeit erreicht für kleine C2X-Ausstattungsgrade einen guten Wert, die Ergebnisse der Schätzung sind dabei jedoch sehr ungenau. Mit steigender C2X-Penetration verbessert sich die Vollständigkeit weiter.

Der Einfluss der Bluetooth-Detektionen kann ebenfalls untersucht werden. Da eine realistische Bluetooth-Ausstattung von 30 - 50% angenommen wurde, wurden Experimente für 30% und 50% durchgeführt. Um den Einfluss des Detektionsverfahrens insgesamt zu beurteilen, wurden auch Ergebnisse für die Extremfälle von 0% und 100% Ausstattungsgrad durchgeführt.

Bei fester C2X-Ausstattung von 10% wird damit das in Tabelle 5.2 dargelegte Ergebnis erreicht.

Es ist zu sehen, dass die Genauigkeit für das Fehlen von mit Bluetooth ausgestatte-

BT-Ausstattung [%]	MAE [m/s]	RMSE [m/s]	Vollständigkeit [%]
0	0,98	0,99	27,42
30	3,35	4,96	61,54
50	3,74	5,37	75,44
100	4,06	5,63	92,74

Tabelle 5.2: Ergebnisse in Abhängigkeit der Bluetooth-Ausstattungsgrade bei fester C2X-Ausstattung von 10%.

ten Fahrzeugen am höchsten ist. Dies ist dadurch zu erklären, dass die C2X-Detektion genauer ist als die Bluetooth-Detektion. So wird für reine C2X-Ausstattung stets ein sehr genaues Ergebnis erzielt. Dabei wird für 10% C2X-Ausstattung und das Fehlen von Bluetooth-Fahrzeugen ein ähnlicher Fehler erreicht wie für eine Ausstattung von 100% C2X-Fahrzeugen mit 30% Bluetooth-Penetration wie in Tabelle 5.1. Ein signifikanter Unterschied besteht jedoch in der Vollständigkeit. Diese liegt für reine C2X-Ausstattung von 10% nur bei etwa 27%. Das Modell liefert folglich sehr exakte Ergebnisse, die aber sehr unvollständig sind. Dies unterstützt die zu Beginn aufgestellte Beobachtung, dass C2X-Daten zwar genau sind, jedoch nicht in ausreichender Zahl für eine Verkehrszustandschätzung vorliegen. Dies hat motiviert, das Bluetooth-Verfahren als zweiten Sensor zur Unterstützung hinzuzunehmen.

Wird die Bluetooth-Penetration erhöht, steigt die Vollständigkeit schnell. Bei 30%-Ausstattung werden schon über 60% erreicht. Gleichzeitig sinkt die Genauigkeit der Schätzung. Besonders im direkten Vergleich der Experimente in denen entweder nur Bluetooth- oder nur C2X-Ausstattung vorliegt, wird deutlich, dass die beiden Sensoren sehr unterschiedliche Ergebnisse liefern: Die C2X-Detektion liefert wesentlich genauere Ergebnisse als die Bluetooth-Detektion. Letztere ist nützlich, da sie mit einem höheren Ausstattungsgrad angenommen werden kann und so die Vollständigkeit der Schätzung signifikant verbessert.

Anhand von Szenario Minimal kann speziell untersucht werden, welchen Einfluss Geschwindigkeit und Verkehrsstärke auf das Ergebnis haben. Dafür sind im Folgenden die Verläufe der tatsächlichen Geschwindigkeit und die Ergebnisse der Geschwindigkeitsschätzung über die Zeit dargestellt.

Die tatsächliche Geschwindigkeit verringert sich dabei dreimal schrittweise von 14 m/s auf 2 m/s. Im ersten Teil wird eine niedrige, im zweiten Teil eine mittlere und im dritten Teil eine hohe Verkehrsstärke von 360, 720 und 1440 Fahrzeugen pro Stunde angesetzt. Es ist zu sehen, dass die Geschwindigkeitsschätzung für niedrige Verkehrsstärken besser funktioniert als für hohe. Insgesamt wird die Geschwindigkeit oft überschätzt, besonders oft wird Geschwindigkeitsklasse 13 geschätzt, was stets vorkommt, wenn Fahrzeuge nur ein einziges Mal mit dem Bluetooth-Sender detektiert werden können. Es ist deutlich zu sehen, dass die Erhöhung der C2X-Ausstattungsrate von 20% (links in 5.5) auf 50%

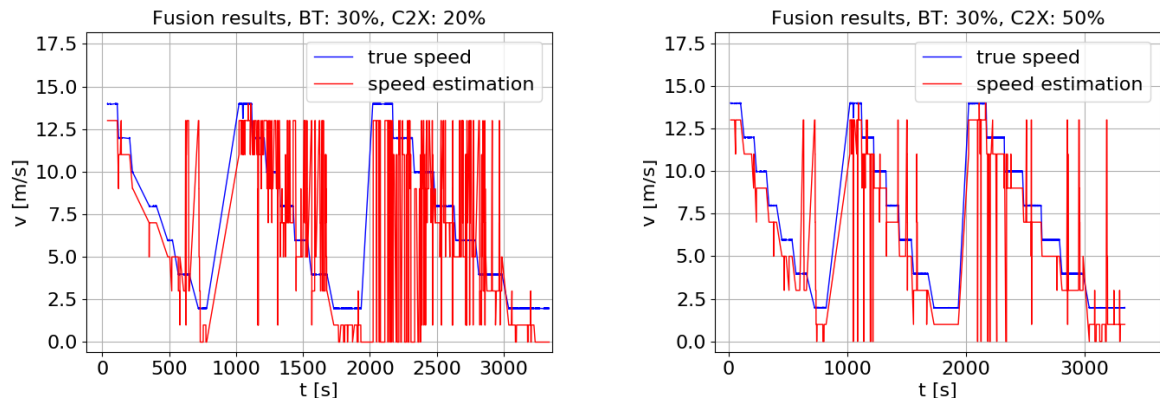


Abbildung 5.5: Vergleich von Schätzung und tatsächlicher Geschwindigkeit für eine Bluetooth-Ausstattung von 30% und C2X-Ausstattungsgrade von 20% (links) und 50% (rechts).

(rechts) die Geschwindigkeitsschätzung verbessert. Wesentlich seltener werden Geschwindigkeiten fundamental falsch eingeschätzt.

Da das Szenario Minimal keine Kreuzung enthält, kann der Knoten *TLC* nicht genutzt werden. Es ist zu beachten, dass im Vergleich mit den folgenden beiden Szenarien ein Einflussfaktor weniger in die Schätzung eingeht.

5.2.2 Szenario RiLSA

Im Szenario RiLSA wird das Modell an einem Knotenpunkt mit minimalistischer Geometrie getestet. Fahrzeuge kommen aus den vier Himmelsrichtungen Norden, Osten, Süden und Westen. Auf jedem Kreuzungsarm sind C2X- und Bluetooth-Detektoren modelliert, auf Grundlage derer jeweils eine Schätzung entsteht. Das Modell liefert insgesamt eine Geschwindigkeitsschätzung pro Kreuzungsarm.

Im Mittelpunkt der Untersuchung des Szenarios steht die Lichtsignalanlage, die als wichtiger Einflussfaktor ins Modell eingeht. Schon bei der Ermittlung der Sensorlikelihoods in den Abbildungen 5.6 und 5.7 ist zu sehen, dass die Phase der Lichtsignalanlage Einfluss auf die Geschwindigkeitsschätzung nimmt. Der erste Graph stellt die Sensorlikelihoods für Phase *rot*, der zweite die für Phase *grün2* dar.

Während für niedrige tatsächliche Geschwindigkeiten Wahrscheinlichkeiten über 60% auftreten, sind die meisten Wahrscheinlichkeiten kleiner, dafür werden mehr Klassen erreicht. Es treten also mehr unterschiedliche Detektionen der gleichen tatsächlichen Geschwindigkeit auf. Insgesamt werden die Geschwindigkeiten eher zu hoch, als zu niedrig

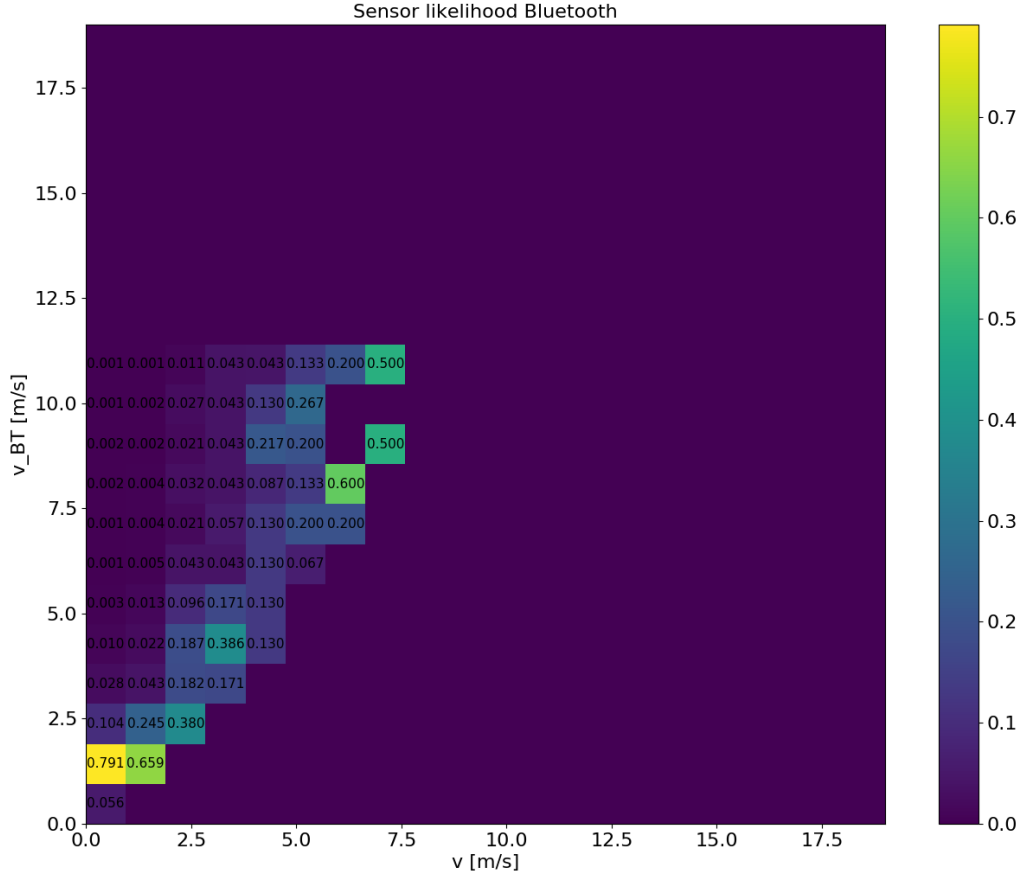


Abbildung 5.6: Sensorlikelihoods für Bluetooth-Detektion für die Phase *rot*.

eingeschätzt. Dies passt zu den allgemeinen Eigenschaften der Bluetooth-Detektion in 2.2.2.

In Abbildung 5.7 wird das noch deutlicher. Während die Detektionen für niedrige tatsächliche Geschwindigkeiten auf viele Klassen verteilt sind, fallen hohe Werte bei 11 m/s für die höheren Geschwindigkeiten auf. Dies kombiniert die allgemeinen Beobachtungen, dass Geschwindigkeiten stets höher geschätzt werden, bei 11 m/s jedoch das absolute Maximum des Detektionsverfahrens erreicht ist. Die erste Matrix gilt für Phase *rot*, die zweite für Phase *grün2*. Während bei *rot* eher niedrigere Geschwindigkeiten auftreten, überwiegen hier höhere Geschwindigkeiten. Auch die Schätzung der Likelihoods folgt diesem Schema. Eine Tendenz zu höheren Schätzungen ist zu erkennen. Die Likelihoods

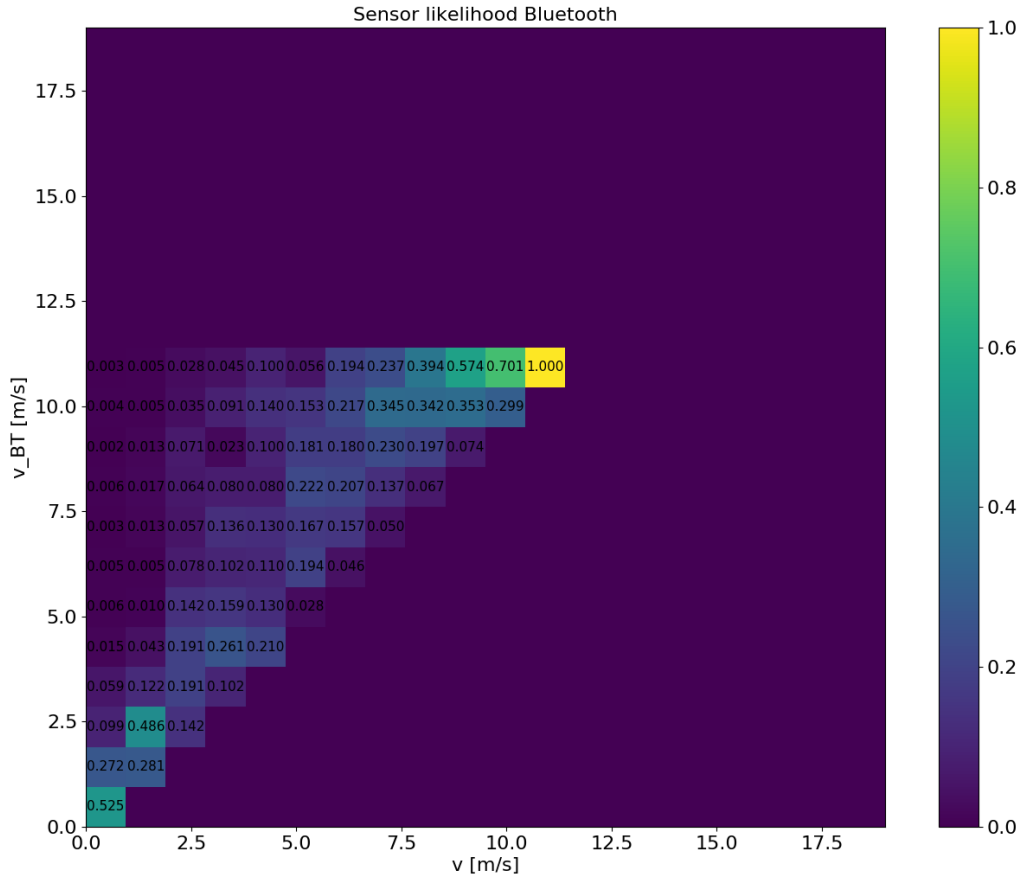


Abbildung 5.7: Sensorlikelihoods für Bluetooth-Detektion für die Phase *grün2*.

für *grün1* sind dabei erwartungsgemäß wertemäßig zwischen den beiden dargestellten Likelihoods zu verorten.

Die C2X-Likelihood unterscheidet sich weniger, wenn unterschiedliche Phasen der Lichtsignalanlage vorliegen. Exemplarisch werden die Sensorlikelihoods für Phase *grün2* in 5.8 dargestellt.

Wie in Szenario Minimal zeichnen sich die C2X-Sensorlikelihoods dadurch aus, dass die höchsten Werte an der jeweiligen tatsächlichen Geschwindigkeitsklasse angenommen werden. Einzige Ausnahme bildet dabei Klasse 0. Mit über 70%-iger Wahrscheinlichkeit werden die Geschwindigkeiten korrekt geschätzt. Doch auch die fehlerhaften Schätzungen

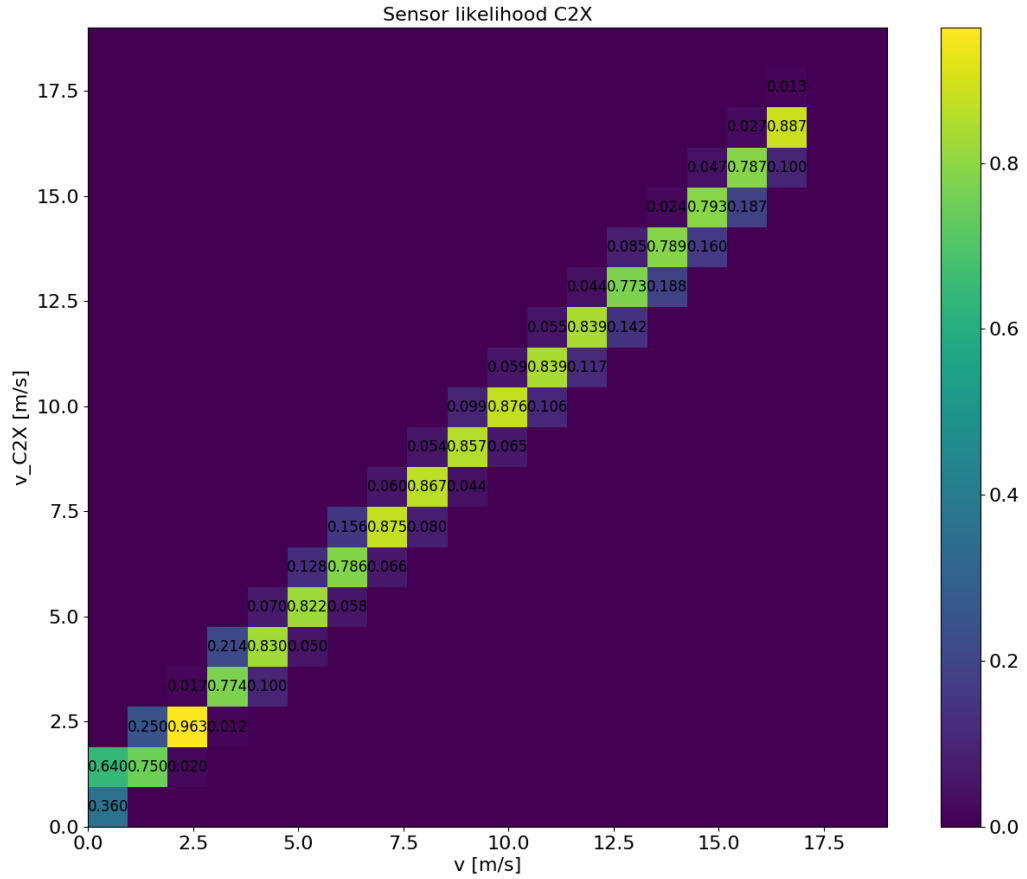


Abbildung 5.8: Sensorlikelihoods für C2X-Detektion für die Phase *grün2*.

erweisen sich als gute Annäherungen: Wird nicht die richtige Klasse geschätzt, so werden zumindest die benachbarten Klassen getroffen. Schätzungen, die nicht in die nächsthöhere oder nächstniedrigere Klasse fallen, treten nicht auf.

Nach den bisherigen Analysen der Lernergebnisse, sollen im Folgenden die Ergebnisse der Datenfusion im Detail betrachtet und analysiert werden. Dabei sollen insbesondere C2X- und Bluetooth-Penetrationsraten auf ihren Einfluss auf die Schätzung untersucht werden. Zunächst sollen die Schätzungen als Mittel der vier Kreuzungsarme betrachtet werden. In Abhängigkeit der C2X-Penetration werden dann die Ergebnisse in Tabelle 5.3 erreicht. Dabei ist ein fester Bluetooth-Ausstattungsgrad von 30% angenommen.

C2X-Ausstattung [%]	MAE [m/s]	RMSE [m/s]	Vollständigkeit [%]
0	1,71	2,68	63,1
1	1,71	2,68	64,31
5	1,72	2,71	68,56
10	1,72	2,72	73,5
20	1,68	2,68	81,07
50	1,50	2,44	93,12
100	1,26	2,01	99,62

Tabelle 5.3: Ergebnisse in Abhängigkeit der C2X-Ausstattungsgrade bei fester Bluetooth-Ausstattung von 30%, Vergleich des mittleren absoluten Fehlers (MAE), des Root Mean Square Errors (RMSE) und der Vollständigkeit.

Es ist zu sehen, dass MAE und RMSE für 0% C2X-Ausstattung schon 1,71 m/s bzw. 2,68 m/s erreichen und diese Werte bis 20% annähernd konstant bleiben. Erst ab einem Ausstattungsgrad von über 20% verbessern sich beide Werte nochmals. Detaillierte Betrachtungen der einzelnen Kreuzungsarme werden mehr Aufschluss über dieses Verhalten bringen.

Die Vollständigkeit liegt für reine Bluetooth-Ausstattung von 30% bei 63% und erhöht sich mit Erhöhung der C2X-Ausstattung fortlaufend. Bei 100% C2X-Ausstattung wird schließlich eine Vollständigkeit von fast 100% erreicht.

Anschließend sollen die Ergebnisse der einzelnen Kreuzungsarme verglichen werden. Zur Übersichtlichkeit wird hier nur eine Metrik verwendet. Dabei wird der MAE gewählt, da höhere und niedrigere Abweichungen gleich stark ins Gewicht fallen sollen. Die Ergebnisse sind in 5.4 dargestellt.

Dabei ist auffällig, dass sich die Schätzungen auf Ost-West-Achse und Nord-Süd-Achse unterscheiden. Während der MAE in Ost-West-Richtung von ca. 2,3 auf 1,5 m/s fällt, liegt er in Nord-Süd-Richtung im Bereich zwischen etwa 1,2 und 1 m/s. Der Fehler ist also niedriger und verändert sich weniger stark. Noch ein weiterer Unterschied fällt auf: Während sich der MAE für die östliche und westliche Richtung Stück für Stück verringert, steigt er in nördlicher und südlicher Richtung bis etwa 20% Ausstattung und fällt erst dann mit weiterer Erhöhung des Ausstattungsgrades. Dies erklärt die zuvor beobachtete

	MAE [m/s]			
C2X [%]	N	O	S	W
0	1,05	2,36	1,09	2,33
1	1,05	2,36	1,09	2,33
5	1,1	2,33	1,14	2,3
10	1,14	2,3	1,16	2,27
20	1,17	2,2	1,16	2,18
50	1,13	1,93	1,08	1,87
100	1,05	1,58	0,97	1,45

Tabelle 5.4: Ergebnisse in Abhängigkeit der Kreuzungsarme und C2X-Ausstattung.

Stagnation des Mittelwerts bis ca. 20% der Richtungen. Um mögliche Ursachen dafür zu finden, weshalb sich der MAE unterschiedlich verhält, sollen die Kreuzungsarme genauer betrachtet werden.

Es ist festzuhalten, dass schon bei der Modellierung des Szenarios Unterschiede vorliegen: Während auf der Ost-West-Achse in jeder Richtung etwa 1000 Fahrzeuge verteilt auf eine Stunde eingesetzt werden, sind es auf der Nord-Süd-Achse nur je 300 Fahrzeuge. Auch die Lichtsignalanlage ist nicht identisch geschaltet: So dauert die *GRÜN*-Phase der Ost-West-Achse 40 s, die der Nord-Süd-Achse hingegen nur 12 s. Dies hat für die Diskretisierung der Phasen in *grün1* und *grün2* nach 7,5 s auch den Effekt, dass Phase *grün2* deutlich seltener erreicht wird und weniger Daten zur Quantifizierung vorliegen.

Wieder soll auch der Einfluss der Bluetooth-Penetrationsrate untersucht werden. Die Ergebnisse der Experimente sind dabei:

BT-Ausstattung [%]	MAE [m/s]	RMSE [m/s]	Vollständigkeit [%]
0	1,71	2,68	33,23
30	1,72	2,72	73,42
50	1,5	2,44	93,12
100	1,26	2,01	99,62

Tabelle 5.5: Ergebnisse in Abhängigkeit der Bluetooth-Ausstattungsgrade bei fester C2X-Ausstattung von 10%.

Wie in Szenario Minimal steigt die Vollständigkeit bei Erhöhung des Bluetooth-Ausstattungsgrades schnell an. Doch auch der MAE verbessert sich mit Erhöhung der Ausstattung: Während sich die Genauigkeit in Szenario Minimal für steigende Bluetooth-Ausstattung verschlechterte, verbessert sie sich hier. Dies ist ein klares Zeichen dafür, dass die Bluetooth-Detektion in diesem weniger extremen Szenario mit geringeren Geschwindigkeiten und Verkehrsstärken wesentlich bessere Ergebnisse liefert. Dies wird auch durch die deutlich niedrigeren Fehler von 1,26 - 1,72 m/s im Gegensatz von bis

zu 4,06 m/s in Szenario Minimal bekräftigt. Dass im Gegensatz zu Szenario Minimal eine deutliche Verbesserung der Genauigkeit erzielt wurde, kann durch den Einfluss der Lichtsignalanlage *TLC* erklärt werden. Schon bei der Modellierung des Bayes'schen Netzes in Kapitel 4 konnte gezeigt werden, dass die Phase der Lichtsignalanlage einen signifikanten Einfluss auf die Geschwindigkeitsverteilung hat. Dies wird durch die Simulationsergebnisse bestätigt.

Dieses Szenario wurde auch in [24] verwendet. Die dort erreichten Ergebnisse sind in Abschnitt 3.5 dargestellt. Da dort mit einer anderen Verkehrsnachfrage gearbeitet wurde, sind die Ergebnisse nur bedingt vergleichbar. Insgesamt scheint die Neumodellierung die Schätzergebnisse jedoch zu verbessern. Es fällt auf, dass vor allem die Vollständigkeit verbessert werden konnte. Besonders bei geringer C2X-Ausstattung tritt dies ein: Bei 1% C2X-Ausstattung werden etwa 64% erreicht, in [24] waren es nur 36%. Der RMSE lag dort weit gestreut zwischen 1,7 und 5,3 m/s und verbesserte sich mit steigender Penetrationsrate auf 1,6 bis 2,1 m/s. Hier erreicht er zunächst bessere Ergebnisse von 2,7 m/s, landet für hohe Penetrationsraten etwas schlechtere Ergebnisse. Der MAE zeigt, dass die Ergebnisse je nach Kreuzungsarm näher beieinander liegen. Die Richtungsabhängigkeit ist hier kleiner als in [24].

Das Phänomen, dass sich auf der Nord-Süd-Achse die Fehler zunächst vergrößern, ist auch in beiden Untersuchungen anzutreffen. Dies legt nahe, dass es sich um ein systematisches Auftreten handelt. Was genau dessen Auslöser ist, bleibt Gegenstand weiterer Untersuchungen.

5.2.3 Szenario BS Forschungskreuzung

Das Szenario BS Forschungskreuzung wird schließlich dafür verwendet, das Modell an einem möglichst realitätsnahen Beispiel zu testen. Sowohl Kreuzungsgeometrie als auch Verkehrsnachfrage sind realen Gegebenheiten nachempfunden. Eine genau vermessene Forschungskreuzung mit zugehörigen Messungen der Verkehrsnachfrage wurden in die Verkehrssimulation übertragen. So können Untersuchungen unter realitätsnahen Bedingungen durchgeführt werden.

Die Ergebnisse des Lernalgorithmus ähneln denen aus Szenario RiLSA. Die Sensorlikelihoods für Bluetooth- und C2X-Detektionen sind dabei ähnlich in ihrer Struktur. Die Schätzergebnisse hingegen sollen im Folgenden gesondert betrachtet werden. Für MAE, RMSE und Vollständigkeit ergeben sich die in Tabelle 5.6 dargestellten Resultate.

Es fällt auf, dass sich bei Erhöhung des C2X-Ausstattungsgrades sowohl Genauigkeit als auch Vollständigkeit verbessern. So verringern sich MAE und RMSE, während die Vollständigkeit sich erhöht. Dies steht im Gegensatz zu den Resultaten in Szenario RiLSA, wo ein anfängliches Stagnieren der Genauigkeit für manche Richtungen beobachtet wurde. Im Folgenden wird diese Auffälligkeit genauer beleuchtet und mögliche Ursachen werden

C2X-Ausstattung [%]	MAE [m/s]	RMSE [m/s]	Vollständigkeit [%]
0	2,37	3,69	64,31
1	2,3	3,62	65,7
5	2,07	3,36	70,23
10	1,86	3,08	74,61
20	1,59	2,71	81,58
50	1,2	2,06	92,84
100	0,98	1,63	99,8

Tabelle 5.6: Ergebnisse in Abhängigkeit des C2X-Ausstattungsgrades bei fester Bluetooth-Ausstattung von 30%.

identifiziert. Zunächst sollen hierfür die Fusionsergebnisse der einzelnen Richtungen betrachtet werden.

Für die Schätzungen auf den Kreuzungsarmen im direkten Vergleich ergibt sich Tabelle 5.7.

C2X [%]	MAE [m/s]			
	N	O	S	W
0	2,32	2,05	2,46	2,64
1	2,28	2,01	2,38	2,53
5	2,02	1,89	2,09	2,29
10	1,8	1,73	1,79	2,09
20	1,54	1,55	1,49	1,76
50	1,18	1,24	1,08	1,3
100	0,98	1,02	0,88	1,04

Tabelle 5.7: Ergebnisse in Abhängigkeit der Kreuzungsarme und C2X-Ausstattung.

Es ist zu sehen, dass sich auch bei Betrachtung der Kreuzungsarme im Einzelnen Genauigkeit und Vollständigkeit verbessern. Die MAEs der einzelnen Richtungen unterscheiden sich um nicht mehr als 0,6 m/s und sind damit ähnlicher als im Szenario RiLSA. Ihre Größenordnung entspricht für niedrige Ausstattungsgrade etwa der Ost-West-Achse im RiLSA-Szenario. Bei höheren Ausstattungsgraden können hier jedoch bessere Ergebnisse erzielt werden.

Die Variation der Bluetooth-Penetration führt zu den in Tabelle 5.8 dargestellten Ergebnissen.

Die Vollständigkeit verändert sich mit steigender Bluetooth-Ausstattung ähnlich wie in den bisher betrachteten Szenarien. Für höhere Penetrationsraten wird eine größere

BT-Ausstattung [%]	MAE [m/s]	RMSE [m/s]	Vollständigkeit [%]
0	1,32	2,22	40,03
30	1,86	3,08	74,61
50	1,94	3,17	84,64
100	2,05	3,29	94,82

Tabelle 5.8: Ergebnisse in Abhängigkeit der Bluetooth-Ausstattungsgrade bei fester C2X-Ausstattung von 10%.

Vollständigkeit erreicht. Die Werte liegen dabei für geringe Ausstattung etwas höher und für hohe Ausstattung etwas niedriger als im Szenario RiLSA. Im Szenario Minimal wurde insgesamt eine niedrigere Vollständigkeit erreicht, was in Zusammenhang mit der insgesamt höheren Geschwindigkeit stehen kann.

Die Genauigkeit hingegen verschlechtert sich mit höherer Bluetooth-Ausstattung. Dies wurde auch in Szenario Minimal beobachtet und liegt darin begründet, dass die Bluetooth-Detektionen naturgemäß weniger exakt sind als die C2X-Detektionen. Bemerkenswert ist jedoch, dass in Szenario RiLSA ein gegenteiliges Ergebnis eintritt: Hier verkleinert sich der Fehler mit höherer Bluetooth-Ausstattung. In diesem Szenario sind noch weitere Auffälligkeiten zu beobachten: Der MAE unterscheidet sich stark zwischen Nord-Süd bzw. Ost-West-Achse. Während letzterer in einem ähnlichen Bereich liegt wie in Szenario BS Forschungskreuzung, ist der MAE in Nord-Süd-Richtung signifikant kleiner. Bei Erhöhung der Penetrationsrate bis 20% wächst er zunächst, bis er bei weiterer Erhöhung wieder auf sein Ausgangsniveau fällt.

Wieso sich in Szenario RiLSA solche Auffälligkeiten zeigen, die in den anderen Szenarien nicht zu finden sind, muss zukünftig noch untersucht werden.

Insgesamt kann also festgestellt werden, dass das Modell zu jedem Zeitschritt Geschwindigkeitsschätzungen liefert, die am Knotenpunkt um nicht mehr als 2,4 m/s von der tatsächlichen Geschwindigkeit abweichen. Damit konnten die Ergebnisse aus [24] besonders für niedrige C2X-Ausstattungsgrade verbessert werden. Im Test ohne Knoten unter atypischen Bedingungen wurden schlechtere Ergebnisse erzielt. Hier konnte die Erkenntnis gewonnen werden, dass sich eine hohe Verkehrsstärke negativ auf die Genauigkeit der Schätzung auswirkt.

Die Vollständigkeit wird maßgeblich vom Bluetooth-Ausstattungsgrad verbessert, während die C2X-Penetrationsrate die Genauigkeit der Schätzung verbessern kann. Dies stimmt mit den Erwartungen an die beiden Detektionsarten überein.

Im Szenario RiLSA konnte auf der Nord-Süd-Achse ein abweichendes Verhalten festgestellt werden, das noch weiterer Klärung bedarf.

6 Fazit

Mit der Entwicklung neuer Technologien verändern sich Chancen und Anforderungen an die Mobilität fortlaufend. Dadurch bieten sich auch neue Möglichkeiten, Verkehr zu erfassen. Kann beispielsweise die Nachrichtenübermittlung per C2X-Kommunikation genutzt werden, um den Verkehr erfassbar zu machen, so kann auf kostenintensive Infrastrukturelemente wie etwa Induktionsschleifen verzichtet werden. Verkehrszustände können auf neue Art und Weise, genauer oder kostengünstiger erfasst werden, als es bisher geschieht. Dafür müssen neue Strategien und Methoden entwickelt werden.

In der vorliegenden Arbeit wurde die Frage untersucht, wie die C2X-Kommunikation genutzt werden kann, um eine Verkehrszustandsschätzung zu ermöglichen. Da die C2X-Technologie noch nicht standardmäßig in Fahrzeugen vorhanden ist, ist davon auszugehen, dass nur ein Teil der Fahrzeuge C2X-Nachrichten senden und empfangen kann. Ein solcher Mischverkehr aus Fahrzeugen mit und ohne C2X-Ausstattung wurde in der Arbeit betrachtet.

Da so nur von dem entsprechend ausgestatteten Anteil der Fahrzeuge Daten erhoben werden können, sind die Daten nur ein kleiner Ausschnitt des tatsächlichen Verkehrsflusses. Um den lückenhaften Datensatz zu verbessern, wurde ein weiterer Sensor zur Detektion verwendet: Mit einem Bluetooth-basierten Verfahren konnten zusätzlich Aussagen über all jene Fahrzeuge getroffen werden, die ein Bluetooth-Gerät mitführen.

Eine Datenfusion bietet die Möglichkeit, die beiden Detektionsarten zu kombinieren und ein gemeinsames Ergebnis zu erhalten. Dabei ist zu beachten, dass sich die Daten in Qualität und Quantität unterscheiden. So sind Geschwindigkeitsdaten aus der C2X-Kommunikation in der Regel genauer als solche, die durch das Bluetooth-Verfahren geschätzt werden. Sie werden mittels GPS aus relativ genauen Positionsdaten bestimmt. Die Bluetooth-Daten hingegen zeichnen sich durch einen größeren Fehler aus, der im Schätzverfahren begründet ist. Bluetooth ist kein Sensor zur Geschwindigkeitsdetektion im eigentlichen Sinne. Das in Abschnitt 2.2.2 vorgestellte Verfahren ermöglicht lediglich eine vage Schätzung der Geschwindigkeit. Dennoch bietet das Bluetooth-Verfahren den Vorteil, dass aufgrund der höheren Ausstattungsrate mehr Fahrzeuge erfasst werden können. Somit liefert das Bluetooth-Verfahren zwar fehlerbehaftetere, dafür mehr Daten als die C2X-Detektion. Diese liefert genauere Daten in geringerer Menge.

C2X- und Bluetooth-Daten werden mit dem Ziel, die Genauigkeit und Vollständigkeit der Schätzung insgesamt zu erhöhen, zu einer gemeinsamen Schätzung fusioniert. Es wurde ein Modell entwickelt, mit dem eine solche Datenfusion ausgeführt werden kann.

Dabei wurde einem probabilistischer Ansatz basierend auf Bayes'schen Netzen gefolgt. Im Modell werden dafür die wichtigsten Verkehrskenngrößen und Einflussfaktoren bestimmt. Die Auswahl orientiert sich dabei auch an der realistischen Einsetzbarkeit.

Anhand von experimentellen Untersuchungen hat sich gezeigt, dass die folgenden Größen geeignet sind, um eine Schätzung des Verkehrszustands zu ermöglichen:

- V : Tatsächliche mittlere Momentangeschwindigkeit,
- V_{BT} : Geschwindigkeitsdetektion mittels Bluetooth-Verfahren,
- V_{C2X} : Geschwindigkeitsdetektion mittels C2X-Verfahren,
- TLC : Phase der Lichtsignalanlage,
- V_{pre} : Letzte Bluetooth-Detektion.

Anschließend wurden die kausalen Zusammenhänge zwischen diesen Größen modelliert und in ein Bayes'sches Netz überführt. Insgesamt ergibt sich dabei das Netz in Abbildung 6.1.

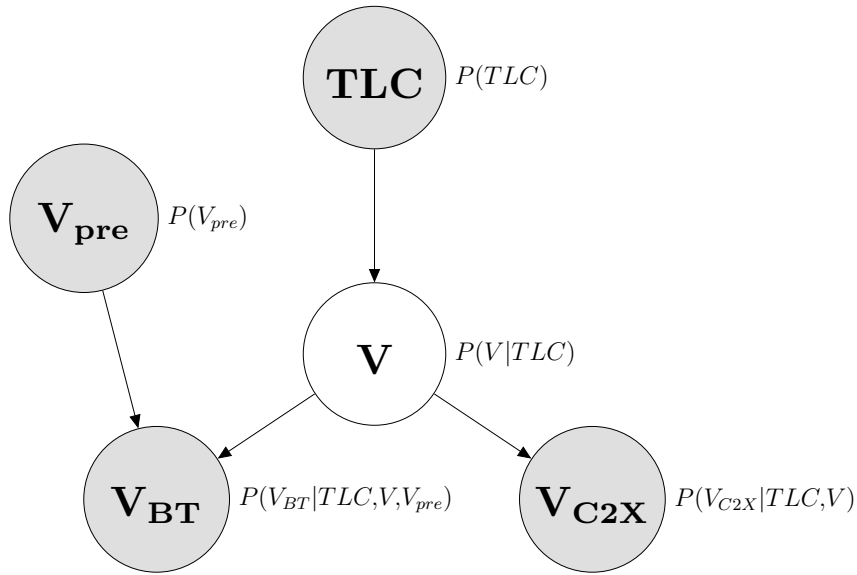


Abbildung 6.1: Struktur des modellierten Bayes'schen Netzes mit bedingten Wahrscheinlichkeiten für alle Knoten.

Bedeutung der Knoten: TLC (traffic light control): Phase der Lichtsignalanlage, V_{pre} : Geschwindigkeit des vorausfahrenden Fahrzeugs, V : Tatsächliche mittlere Momentangeschwindigkeit, V_{BT} : Bluetooth-basierte Geschwindigkeitsschätzung, V_{C2X} : C2X-basierte Geschwindigkeitsschätzung.

Dann wurden die Wahrscheinlichkeitsverteilungen des zugrundeliegenden Bayes'schen Netzes experimentell bestimmt. Hierfür wurde das Modell implementiert und ein Skript

entwickelt, das das Modell in Simulationen in SUMO einsetzbar macht. Aus dem Skript wird der Verkehrssimulator Schritt für Schritt gesteuert und die Detektionsdaten zur Simulationszeit erfasst und direkt verarbeitet. In drei Schritten konnten so Detektionen aufgezeichnet, relative Häufigkeiten zur Bestimmung der Wahrscheinlichkeitsverteilungen berechnet und schließlich die Datenfusion ausgeführt werden.

Für die simulativen Experimente wurden drei Szenarien erstellt, mit Hilfe derer das Modell getestet werden kann:

- Minimal,
- BS Forschungskreuzung,
- RiLSA.

Dabei unterscheiden sich die Szenarien in Anwendungsbereich, Abstraktionslevel und Realitätsnähe.

Das erste Szenario besteht aus einer geraden Straße ohne Knoten. Der Verkehrsfluss wird in festgelegten Zeitintervallen verändert, sodass unterschiedliche Geschwindigkeiten und Verkehrsstärken untersucht werden können. Solche Verkehrsflüsse sind wenig realistisch, ermöglichen jedoch, das Modell unter klar definierten Bedingungen zu untersuchen.

Das zweite Szenario RiLSA ermöglicht die Untersuchung des Modells an einem stark vereinfachten Knotenpunkt. Dabei wird eine minimalistische Kreuzung verwendet, an der vier Kreuzungsarme rechtwinklig aufeinander treffen. Die Geometrien sind dabei simpel gehalten: Alle Kreuzungsarme sind gerade Strecken. So kann das Modell zunächst an einem Knotenpunkt untersucht werden, ohne eine starke Beeinflussung durch komplexe Geometrien zu erfahren.

Schließlich wird im dritten Szenario BS Forschungskreuzung das Modell an einem realitätsnahen Knoten getestet. Die Geometrie des Szenarios ist einer realen Kreuzung nachempfunden. Auch die Verkehrsnachfrage stammt aus einer Messung des Verkehrs an dieser Kreuzung und wurde so realistisch quantifiziert. Dieses Szenario dient der Untersuchung des Modells anhand eines möglichst realitätsnahen Beispiels. In Abbildung 2.6 sind die reale und die modellierte Kreuzung zu sehen.

Die Ergebnisse der Datenfusion zeigen dabei, dass das Modell auf allen Szenarien für C2X-Ausstattungsgraden ab 10 – 20% Ergebnisse liefert, die hinreichend genau und vollständig sind. Während die C2X-Ausstattung maßgeblich zur Verbesserung der Genauigkeit des Modells beiträgt, erhöht die Bluetooth-Penetration die Vollständigkeit der Ergebnisse. Die Datenfusion ermöglicht, auf beiden Qualitätsskalen gute Ergebnisse zu erzielen. Im Vergleich der Szenarien untereinander fällt auf, dass die Schätzung an Knotenpunkten und bei niedriger Verkehrsstärke die besten Ergebnisse liefert. So werden im realitätsnahen Szenario BS Forschungskreuzung für 10% C2X- und 30% Bluetooth-Penetration mittlere absolute Fehler (MAE) von 1,7 – 2 m/s erreicht. Im Szenario RiLSA sind die Ergebnisse teilweise noch besser. Die MAEs liegen im Bereich von 1,1 – 2,3 m/s.

Das Szenario Minimal ist als Test unter extremen Bedingungen zu verstehen: Geschwindigkeit und Verkehrsstärke durchlaufen künstlich den gesamten Wertebereich. Dadurch werden Ergebnisse mit einem MAE von 3,35 m/s erzielt. Die Vollständigkeit dabei für Szenario Minimal über 60% sowie für die Szenarien RiLSA und BS Forschungskreuzung über 70%.

Insgesamt ist festzuhalten, dass das Modell in der Lage ist, eine gute Schätzung des Verkehrszustands zu liefern. Ist der Verkehrszustand erfassbar, kann er zur effizienten Verkehrssteuerung eingesetzt werden. Es kann untersucht werden, ob die Geschwindigkeitsschätzung zur Steuerung einsetzbar ist. Eine Möglichkeit zur Verkehrssteuerung bietet die Lichtsignalanlage. Es kann simulativ erforscht werden, ob und wie die Steuerung der Lichtsignalanlage an die Geschwindigkeitsschätzung des Bayes'schen Netzes gekoppelt werden kann.

Gegenüber dem in [24] beschriebenen Netz konnte eine deutliche Verbesserung von Genauigkeit und Vollständigkeit der Ergebnisse erzielt werden. Da die Ergebnisse je nach Szenario deutlich voneinander abweichen, scheint insbesondere Optimierungspotential in der Anpassung an ein spezielles Szenario zu bestehen. Es bleibt zu untersuchen, ob eine Veränderung des Bayes'schen Netzes oder seiner Parametrisierungen für spezielle Szenarien bessere Ergebnisse liefern können. Besonders das Szenario RiLSA wirft noch Fragen auf. Auf der wenig frequentierten Nord-Süd-Achse sind die Schätzergebnisse bei sehr niedrigen C2X-Ausstattungsgraden besonders gut, verschlechtern sich aber bei Erhöhung auf bis zu 20%. Dieses Verhalten steht möglicherweise in Zusammenhang mit der kurzen *GRÜN*-Phase der Lichtsignalanlage oder einer niedrigen Verkehrsstärke auf der Achse. Die tatsächliche Ursache kann jedoch nur durch weitere Untersuchungen identifiziert werden.

Auch wenn eine Lichtsignalanlage vorhanden ist, können deren Phasen noch weiter untersucht werden. Insbesondere ist zu klären, wie mit den Phasen *GELB* und *ROT/GELB* zu verfahren ist. Die Vereinfachung, diese Phasen zu Phase *GRÜN* zu zählen, ist im Verkehrssimulator plausibel, sollte aber bei praktischer Anwendung kritisch hinterfragt werden.

Weiter kann untersucht werden, wie sich das Modell auch auf andere Bereiche als innerstädtische Knoten ausweiten lässt. Während hier die Phase der Lichtsignalanlage als wichtiger Faktor identifiziert wurde, steht er in anderen Verkehrssituationen möglicherweise nicht zur Verfügung. Eine geeignete Modifikation des Netzes könnte jedoch auch die Anwendbarkeit in anderen Situationen ermöglichen. Inwieweit das Modell den Verkehr auch dort treffend abbilden kann, muss gesondert untersucht werden.

Literaturverzeichnis

- [1] ABI research: *Global V2V Penetration into New Vehicles to Rise to 70% in 2027, Driven by Both Mandates and Commercial Consortiums*, unter: <https://www.abiresearch.com/press/global-v2v-penetration-into-new-vehicles-to-rise-t/> (abgerufen am 08.11.2017).
- [2] R. Baldessari, B. Bödekker, A. Brakemeier, M. Deegener, A. Festag, W. Franz, A. Hiller, C. Kellum, T. Kosch, A. Kovacs, M. Lenardi, A. Lübke, C. Menig, T. Peichl, M. Roeckl, D. Seeberger, M. Strassberger, H. Stratil, H.-J. Vögel, B. Weyl, W. Zhang: *Car-2-car communication consortium manifesto*. Car 2 Car Communication Consortium, 2007.
- [3] L. Bieker, D. Krajzewicz, M. Röck, H. Cappelle: Derivation of a fast, approximating 802.11p simulation model, in *Intelligent Transport Systems Telecommunications*, 2010.
- [4] O. Brandl: V2X traffic management, in *Elektrotechnik & Informationstechnik*, pp. 335 -355, 2016.
- [5] G. Casella, R. L. Berger: *Statistical Inference*. Duxbury Press, 2002.
- [6] T. Chai, R. R. Draxler: Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature, in *Geoscientific Model Development*, pp. 1247-1250, 2014.
- [7] Y.-J. Choi, J. Hur, H.Y. Jeong, C. Joo: Special issue on V2X communications and networks, in *Journal of Communications and Networks*, pp. 205 - 208, 2017.
- [8] I.-W. Chomjakov: *Bluetooth: So klappt jede Verbindung*, unter: <https://www.pcwelt.de/a/geraete-ueber-bluetooth-verbinden,3051531> (abgerufen am 09.11.2017).
- [9] P. Congdon: *Bayesian Statistical Modelling*. Wiley, 2007.
- [10] K. Dresner: *autonomous intersection management*, unter: <http://www.cs.utexas.edu/~aim/> (abgerufen am 08.11.2017).
- [11] S. Even: *Graph Algorithms*. Cambridge University Press, 2011.

- [12] Forschungsgesellschaft für Straßen- und Verkehrswesen, Arbeitsgruppe Bemessung von Straßenverkehrsanlagen: *Handbuch für die Bemessung von Straßenverkehrsanlagen*. FGSV-Verlag, 2015.
- [13] Forschungsgesellschaft für Straßen- und Verkehrswesen, Arbeitsgruppe Verkehrsführung und Verkehrssicherheit: *Richtlinien für Lichtsignalanlagen: RiLSA ; Lichtzeichenanlagen für den Straßenverkehr*. FGSV-Verlag, 2015.
- [14] Forschungsgesellschaft für Straßen- und Verkehrswesen, Arbeitsgruppe Verkehrsführung und Verkehrssicherheit: *Beispielsammlung zu den Richtlinien für Lichtsignalanlagen*. FGSV-Verlag, 2010.
- [15] C. Fox, A. Levitin, T. Redman: The notion of data and its quality dimensions, in *Information Processing & Management*, pp. 9 - 19, 1994.
- [16] Garmin: *What is GPS?*, unter: <http://www8.garmin.com/aboutGPS/> (abgerufen am 08.11.2017)
- [17] T. Gerke, M. Roggero: *Auf dem Weg zum autonomen Fahren*, unter: <http://www.all-electronics.de/auf-dem-weg-zum-autonomen-fahren/> (abgerufen am 12.12.2017).
- [18] Gps.gov.: *GPS Accuracy*, unter: <https://www.gps.gov/systems/gps/performance/accuracy/> (abgerufen am 06.11.2017).
- [19] D. L. Hall, J. Llinas: An Introduction to Multisensor Data Fusion, in *Proceedings of the IEEE* , pp. 6-23, 1997.
- [20] D. Heckerman: *A Tutorial on Learning With Bayesian Networks*. Technical Report MSRTR-95-06, <http://research.microsoft.com/en-us/um/people/heckerman/tutorial.pdf>, Microsoft Research, 1996.
- [21] heise Autos: *Die ersten C2X-Autos kommen auf die Straße. Gute Unterhaltung*, unter: <https://www.heise.de/autos/artikel/Die-ersten-C2X-Autos-kommen-auf-die-Strasse-3025091.html> (abgerufen am 08.11.2017).
- [22] C. Holzmann, S. Oppl: *Bluetooth in a Nutshell*. JKU Linz, 2003.
- [23] V. Joshi, N. Rajamani, K. Takayuki, N. Prathapaneni, L. V. Subramaniam: Information Fusion Based Learning for Frugal Traffic State Sensing, in *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, pp. 2826-2832, 2013.
- [24] M. Junghans, A. Leich: *Traffic State Estimation with Bayesian Networks at Extremely Low V2X Penetration Rates*. 19th International Conference on Information Fusion, Juli 2016.

- [25] B. Kipper, T. Brants, M. Plach, R. Schäfer: *Bayessche Netze: Ein einführendes Beispiel*. Schriftenreihe des Graduiertenkollegs Kognitionswissenschaft, Bericht Nr. 4, 1995.
- [26] U. B. Kjaerulff, A. L. Madsen: *Bayesian Networks and Influence Diagrams*. Springer Science+ Business Media, 2008.
- [27] L. A. Klein: *Sensor and Data Fusion Concepts and Applications*. SPIE Press, 1999.
- [28] K. R. Koch: *Einführung in die Bayes-Statistik*. Springer, 2000.
- [29] D. Krajczewicz, J. Erdmann, M. Behrisch, L. Bieker: Recent Development and Applications of SUMO – Simulation of Urban MObility, in *International Journal On Advances in Systems and Measurements*, pp. 128-138, 2012.
- [30] S. Krauß: *Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics*. Dissertation, Universität zu Köln, 1998.
- [31] C. Lochert, B. Scheuermann, M. Caliskan, M. Mauve: *The Feasibility of Information Dissemination in Vehicular Ad-Hoc Networks*. WONS 2007, 4. Conference on Wireless On demand Network Systems and Services, Januar 2007.
- [32] M. Maurer, J. C. Gerdes, B. Lenz, H. Winner: *Autonomes Fahren Technische, rechtliche und gesellschaftliche Aspekte*. Springer, 2015.
- [33] Mercedes Benz: *Car-to-X communication*, unter: <https://www.mercedes-benz.com/en/mercedes-benz/innovation/car-to-x-communication/> (abgerufen am 08.11.2017)
- [34] R. E. Neapolitan: Is Higher-Order Uncertainty Needed?, in *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, Vol. 26, No. 3, 1996.
- [35] R. E. Neapolitan: *Learning Bayesian Networks*. Pearson Education, Inc., Prentice-Hall, 2003.
- [36] R. Oertel: *Verlustzeitbasierte Steuerung von Lichtsignalanlagen*. Dissertation, TU Dresden, 2014.
- [37] T. Otterstätter: *Methoden zur Erfassung von Verkehrsströmen und Fahrzeiten mit stationären fahrzeugwiedererkennenden Detektoren*. Dissertation, Universität Stuttgart, 2013.
- [38] B. S. Peterson, R. O. Baldwin, J. P. Kharoufeh: Bluetooth Inquiry Time Characterization and Selection, in *IEEE Transactions on Mobile Computing*, pp. 1173-1187, 2006.

- [39] S. von der Ruhren: *Kurzfristprognosen von Verkehrszuständen auf Basis von Verfahren der Mustererkennung und von dynamischen Routensuch- und Umlegungsverfahren*. Dissertation, RWTH Aachen, 2006.
- [40] S. Ruppe, M. Junghans, M. Haberjahn, C. Troppenz: *Augmenting the Floating Car Data Approach by Dynamic Indirect Traffic Detection*. Transport Research Arena - Europe, 2012.
- [41] W. Schnabel, D. Lohse: *Grundlagen der Straßenverkehrstechnik und der Verkehrsplanung, Band 1: Straßenverkehrstechnik*. Beuth Verlag, 2011.
- [42] M. Schulze: *DRIVE C2X - Project Overview*, unter: <http://www.drive-c2x.eu/publications> (abgerufen am 06.10.2017), Daimler AG, 2014.
- [43] Senatsverwaltung für Umwelt, Verkehr und Klimaschutz: *Ampeln und Co., Grundlagen*, unter: <http://www.berlin.de/senuvk/verkehr/lenkung/ampeln/grundlagen/index.shtml> (abgerufen am 16.08.2017).
- [44] M. Treiber, A. Kesting: *Verkehrsdynamik und -simulation*. Springer, 2010.
- [45] P. Wagner: Steuerung und Management in einem Verkehrssystem mit autonomen Fahrzeugen, in *Autonomes Fahren, Technische, rechtliche und gesellschaftliche Aspekte*. Springer, 2015.
- [46] A. Wegener, M. Piorkowski, M. Raya, H. Hellbrück, S. Fischer, J.-P. Hubaux: TraCI: An Interface for Coupling Road Traffic and Network Simulators, in *11th Communications and Networking Simulation Symposium (CNS)*, pp. 155-163, 2008.
- [47] L. Wischoff: *Self-Organizing Communication in Vehicular Ad Hoc Networks*. Dissertation, Technische Universität Hamburg-Harburg, 2007.
- [48] *Definition of Vehicles, Vehicle Types, and Routes*, unter: http://sumo.dlr.de/wiki/Definition_of_Vehicles,_Vehicle_Types,_and_Routes (abgerufen am 10.10.2017).
- [49] *TRAFFIC PARAMETER DEFINITIONS*, unter: https://ntl.bts.gov/DOCS/96100/ch02/ch02_03.html (abgerufen am 18.12.2017).
- [50] *Images of the XCYCLE Stationary demonstrator in Braunschweig*, unter: <http://www.xcycle-h2020.eu/Development/CooperativeSystem> (abgerufen am 10.12.2017).
- [51] *Sensor Fusion Algorithm Development*, unter: <https://www.baselabs.de/create/> (abgerufen am 18.12.2017).